

UNIVERSIDAD COMPLUTENSE DE MADRID

**FACULTAD DE CIENCIAS QUÍMICAS
DEPARTAMENTO DE QUÍMICA ANALÍTICA**



TESIS DOCTORAL

**Desarrollo y aplicación de métodos quimiométricos
para el estudio de muestras mediante
Espectroscopia de Ablación Láser (LIBS)**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Samuel Moncayo Martín

DIRECTOR

Jorge Omar Cáceres Gianni

Madrid, 2017



UNIVERSIDAD
COMPLUTENSE
MADRID

Samuel Moncayo Martín

**Desarrollo y Aplicación de Métodos Quimiométricos para
el Estudio de Muestras Mediante Espectroscopia de
Ablación Laser (LIBS)**

Tesis Doctoral

Departamento de Química Analítica

Madrid 2016



Samuel Moncayo Martín

Certificado

Jorge Omar Cáceres Gianni, Doctor en Ciencias Químicas, Profesor Titular del Departamento de Química Analítica de la Facultad de Ciencias Químicas de la Universidad Complutense de Madrid.

CERTIFICA:

Que el presente trabajo titulado **DESARROLLO Y APLICACIÓN DE MÉTODOS QUIMIOMÉTRICOS PARA EL ESTUDIO DE MUESTRAS MEDIANTE ESPECTROSCOPIA DE ABLACIÓN LASER (LIBS)** ha sido realizado por el licenciado Samuel Moncayo Martín en el laboratorio de Química Láser del Departamento de Química Analítica de la Universidad Complutense de Madrid bajo mi supervisión y que las publicaciones que lo avalan no han sido utilizadas en tesis anteriores, reuniendo a mi juicio los requisitos necesarios para optar al grado de Doctor por la Universidad Complutense de Madrid, por lo que autorizo su presentación.

Jorge Omar Cáceres Gianni

Doctor en Ciencias Químicas

Universidad Complutense

Madrid, 15 Enero de 2016



Agradecimientos

Esta Tesis doctoral no hubiese sido posible sin el apoyo y la colaboración de innumerables personas y entidades a las que me gustaría conceder mi agradecimiento. He aquí las que permanecerán en mi recuerdo por su esfuerzo y dedicación:

En primer lugar, agradecer a las instituciones que han permitido el desarrollo de esta tesis, fundamentalmente a la Universidad Complutense de Madrid, además de a la Universidad Autónoma de Madrid, el Consejo Superior de Investigaciones Científicas, a la Universidad Comenius de Bratislava (Eslovaquia) y al L'Institut Lumière Matière de la universidad Claude Bernard de Lyon (Francia). Un agradecimiento especial a mi director de Tesis Jorge Cáceres por todo su apoyo, orientación, enseñanza y valores, así como a los demás profesores del grupo de investigación, Roberto Izquierdo, Daniel Rosales y Javier Villena. Sin olvidar al resto de profesores, y particularmente a Fernando Navarro por su paciencia y ayuda en aspectos quimiométricos. Al profesor Joaquín Camacho por ofrecernos su equipo, a Luis Díaz por todo su trabajo en nuestros experimentos, sin olvidar a mi director de estancia el Profesor Pavel Veis de la Universidad Comenius por acogerme y enseñarme todo lo relacionado con los parámetros físicos del plasma así como a todo su grupo, singularmente al Doctor Peter Cermak y los licenciados Jakub Ulik, Mária Suchoňová y Marek Pribula. Al profesor Vincent Motto-ros y al doctor Florian Trichar por su ayuda en la realización de las imágenes elementales.

A mis maravillosas compañeras de laboratorio (Sadía Manzoor, Alicia Marín y Lydia Ugena) por todos los buenos y no tan buenos momentos que hemos pasado en el laboratorio y hacer de esta tesis una etapa importantísima de mi vida no solo a nivel académico sino también personal. Sin desatender a todos los estudiantes de Grado y



Samuel Moncayo Martín

Máster que han colaborado en la realización de algunos experimentos, destacando a Esther Sánchez, Eirini, Elena y sobre todo a Ana Rodrigo.

A todos los autores de las publicaciones originadas, principalmente a Asier García, al Dr. Ismael Coronado y Juncal Cruz .

Este proyecto es también parte de mi familia. Agradezco a mis padres María Victoria y Desiderio por su apoyo, cariño, paciencia y financiación; a mi hermano Mario; y a todos mis abuelos.

Por último y no menos importante, a mis amigos de la facultad: Guillermo, Héctor, María y muchos más, por los buenos ratos pasados y a Elena González por su apoyo y ánimos.



Samuel Moncayo Martín

A mis cuatro abuelos, singularmente a Fernando Moncayo por despertar mi pasión por la ciencia y enseñarme a ser mejor persona. Allá donde esté se sentirá orgulloso de mí.



ÍNDICE:

Resumen:	IX
Objetivos:	XVIII
Publicaciones:	XIX

Capítulo 1: *Introducción General*

1.1 ESPECTROSCOPIA DE ABLACIÓN LÁSER (LIBS)	3
1.2 CARACTERÍSTICAS DE LA TÉCNICA LIBS	5
1.3 FUNDAMENTOS DE LA TÉCNICA LIBS	7
1.3.1 Proceso de ablación.....	9
1.3.2 Ablación con pulsos cortos.....	10
1.3.3 Ablación con pulsos ultracortos.....	11
1.3.4 Mecanismos de ablación.....	12
1.4 EVOLUCIÓN DEL PLASMA INDUCIDO POR LÁSER.....	15
1.5 APLICACIONES LIBS	17

Capítulo 2: *Aspectos analíticos de LIBS: Análisis cualitativo y cuantitativo*

2.1 ANÁLISIS CUALITATIVO MEDIANTE LIBS	19
2.2 ANÁLISIS CUANTITATIVO MEDIANTE LIBS	22
2.3 ENFOQUES PARA REALIZAR UN ANÁLISIS CUANTITATIVO	25
2.3.1 Calibración con estándar interno.....	26
2.3.2 Calibración basada en parámetros físicos del plasma (Calibration Free).....	28
2.4 FIGURAS DE MÉRITO DE LA CLASIFICACIÓN	34
2.5 FIGURAS DE MÉRITO DE LA CUANTIFICACIÓN	36
2.5.1 Precisión	37
2.5.2 Límite de detección y cuantificación	38
2.5.3 Exactitud	38

Capítulo 3: *Métodos Quimiométricos aplicados al análisis LIBS*

3.1 INTRODUCCIÓN A LA QUIMIOMETRÍA	40
3.2 CLASIFICACIÓN DE LOS MODELOS QUIMIOMÉTRICOS	41
3.3 ANÁLISIS EXPLORATORIO DE DATOS (CLASIFICACIÓN NO SUPERVISADA)	45
3.3.1 Análisis de componentes principales (PCA).....	46
3.4 MÉTODOS DE CLASIFICACIÓN SUPERVISADOS LINEALES	49
3.4.1 Análisis discriminante.....	49
3.4.2 Modelado suave independiente por analogía de clases (SIMCA).....	51
3.4.3 Regresión Parcial de Mínimos Cuadrados – Análisis Discriminante (PLS-DA).....	54
3.4.4 Árbol de decisión (CART).....	58
3.4.5 Regresión Logística Binaria (BLR)	59
3.5 MÉTODOS DE CLASIFICACIÓN SUPERVISADOS NO LINEALES.....	60
3.5.1 Máquinas de Vectores Soporte (SVM).....	60
3.6 REDES NEURONALES	62
3.6.1 Introducción a las redes neuronales.....	62
3.6.2 Tipos de Neuronas	64



3.6.3 Entrenamiento de una red neuronal	65
3.6.4 Tipos de asociación entre la información de entrada y de salida	67
3.6.5 Principales tipos de arquitecturas de las redes neuronales	68
3.6.6 Algoritmos de entrenamiento de redes neuronales	72
3.6.7 Generalización	78
3.6.8 Limitaciones de los modelos de Red Neuronal	79

Capítulo 4: Instrumentación y montajes experimentales

4.1 INSTRUMENTACIÓN LIBS UTILIZADA	81
4.1.1 Láseres	81
4.1.2 Posicionamiento de la muestra	84
4.1.3 Sistemas colectores de la radiación empleados	85
4.1.3 Analizadores de la radiación utilizados (espectrómetros)	86
4.1.4 Detectores de la radiación	90
4.2 SISTEMAS EXPERIMENTALES UTILIZADOS	92
4.2.1 Sistema Experimental utilizado en la UCM	92
4.2.2 Sistema experimental utilizado en la UAM	94
4.2.2 Sistema experimental utilizado en la Universidad Comenius (Eslovaquia)	96

Capítulo 5: LIBS y métodos quimiométricos aplicados al análisis alimentario

5.1 ANÁLISIS DE ACEITES DE OLIVA MEDIANTE LIBS Y REDES NEURONALES ARTIFICIALES	99
5.1.1 Introducción	100
5.1.2 Procedimiento Experimental	102
5.1.3 Resultados	108
5.2 LIBS APLICADO AL CONTROL DE FRAUDE EN VINOS	119
5.2.1 Introducción	119
5.2.2 Sistema experimental	122
5.2.3 Principales conclusiones de la identificación de vinos	133

Capítulo 6: LIBS aplicado a la ciencia forense

6. 1 DISCRIMINACIÓN DE CUERPOS HUMANOS EN BASE A HUESOS Y DIENTES POR ESPECTROSCOPIA DE ABLACIÓN LÁSER (LIBS) Y REDES NEURONALES (NN)	138
6.1.1 Introducción a la Identificación de restos óseos mediante LIBS y NN	138
6.1.2 Procedimiento Experimental	140
6.1.3 Resultados	144
6.1.4 Conclusiones a la discriminación de individuos mediante LIBS y NN	152
6. 2 EVALUACIÓN DE MÉTODOS QUIMIOMÉTRICOS SUPERVISADOS PARA LA CLASIFICACIÓN DE RESTOS ÓSEOS HUMANOS MEDIANTE ESPECTROSCOPIA DE ABLACIÓN LASER (LIBS)	153
6.2.1 Introducción a la comparativa de métodos quimiométricos para la discriminación y agrupamiento de restos óseos	154
6.2.2 Métodos aplicados	157
6.2.3 Procedimiento experimental	157
6.2.4 Resultados y discusión	162
6.2.5 Conclusiones a la evaluación de métodos quimiométricos supervisados en la discriminación de individuos mediante LIBS	174



Capítulo 7: LIBS aplicado al análisis Medioambiental

7.1 MEDIDA DE LA RELACIÓN MG/CA POR LIBS: UN NUEVO ENFOQUE PARA DESCIFRAR LAS CONDICIONES AMBIENTALES.....	185
7.1.1 Introducción al análisis de muestras de moluscos	186
7.1.2 Materiales y métodos.....	189
7.1.3 Resultados y discusión.....	193
7.1.4 Conclusiones al análisis de muestras de molusco mediante LIBS	200
7.2 RESULTADOS PRELIMINARES DEL ESTUDIO DE RELACIONES ELEMENTALES POR LIBS: ANÁLISIS DE CORALES	202
7.2.1 Introducción al análisis de corales.....	202
7.2.2 Materiales y métodos.....	204
7.2.3 Resultados y discusión.....	208
7.2.4 Conclusiones al análisis de corales.....	212
 CONCLUSIONS:	 216
 BIBLIOGRAFÍA:	 220
 ANEXOS:	 247



Resumen

La espectroscopía de ablación láser (LIBS) es una técnica analítica que se basa en la formación y estudio de un plasma inducido por un pulso láser focalizado sobre la superficie de un material. La emisión producida por dicho plasma, debido a la recombinación electrónica, es colectada generando un espectro-que permite el estudio cualitativo y/o cuantitativo del material.

LIBS se ha convertido en una técnica espectroscópica establecida con un enorme valor analítico debido al conjunto de ventajas que presenta y sus particulares características técnicas. Es destacable su rapidez, versatilidad y capacidad de producir información analítica para prácticamente cualquier tipo de muestra de forma escasamente destructiva y con mínima o incluso nula preparación de muestra. Su instrumentación es sencilla, robusta y compacta, permitiendo incluso llevar análisis in situ, así como realizar análisis remotos.

El principal objetivo de esta tesis es el estudio de la técnica LIBS en combinación con métodos quimiométricos de análisis, especialmente redes neuronales artificiales, para llevar a cabo estudios cualitativos (identificación y discriminación) usando espectros LIBS como huellas espectrales digitales características de muestras de origen alimenticio (aceites de oliva y vinos), con interés forense (huesos y dientes humanos) y por último especies animales, moluscos y corales, con interés paleoclimático. La técnica LIBS es ampliamente conocida, así como los métodos quimiométricos; sin embargo la combinación de ambos está aún en desarrollo y son pocos los autores que usan métodos quimiométricos avanzados para la mejora del rendimiento de la técnica. Aquí, se ha realizado una importante aportación al estudio e implementación de varios métodos quimiométricos para la realización de análisis



La estructura de la tesis consta de dos partes fundamentales:

- X



discriminante lineal (LDA), modelado suave independiente por analogía de clases (SIMCA), regresión parcial de mínimos cuadrados – análisis discriminante (PLS-DA), árboles de decisión (CART), regresión logística binaria (BLR). Por otra parte se han aplicado métodos no lineales como las redes neuronales artificiales (NN) y las máquinas de soporte vectorial (SVM). Debido al peso importante que tienen las redes neuronales artificiales en el desarrollo de la tesis, en este capítulo se desarrolla un último epígrafe dedicado al estudio de su arquitectura, funciones y algoritmos de entrenamientos más habituales. Con este apartado se pretende dar una visión diferente a las redes neuronales, detallando sus operaciones y así eliminar su imagen de “caja negra”.

- En la segunda parte (Cap. 4 - 7) se presentan los resultados experimentales obtenidos en el desarrollo y aplicación de la técnica LIBS con la intención de cumplir tres objetivos fundamentales.
 - En primer lugar, el capítulo 4 “*Instrumentación y montajes experimentales*” se muestran los diferentes sistemas experimentales utilizados, así como toda la instrumentación.
 - El capítulo 5 “*LIBS y métodos quimiométricos aplicados al análisis alimentario*” aborda el desarrollo de una metodología basada en LIBS y métodos quimiométricos para la detección de la adulteración de aceites de oliva, así como de asegurar y controlar la denominación de origen certificada de vinos españoles. Dado que la combinación entre LIBS con diversos métodos quimiométricos ha despertado cada vez más interés en la comunidad LIBS, se ha planteado un sistema de evaluación de



diferentes métodos quimiométricos para la clasificación de muestras basado en medidas LIBS.

- En el capítulo 6 “*LIBS aplicado a la ciencia forense*” hemos estudiado y comparado los rendimientos de siete métodos quimiométricos en la clasificación de muestras de interés forense (huesos y dientes humanos) a través de espectros LIBS.
- Por último, en el capítulo 7 “*LIBS aplicado al análisis Medioambiental*” se presentan los resultados de los estudios desarrollados con muestras de interés medioambiental y arqueológico. Se ha medido, por primera vez, la relación Mg/Ca en conchas de moluscos mediante LIBS y los resultados obtenidos se han comparado con los valores de las medidas de isótopos estables de oxígeno, para su correlación con la temperatura superficial del agua del mar y una futura correlación con datos del pasado (periodo mesolítico). Se presentan también los resultados experimentales preliminares del estudio medioambiental realizado sobre esqueletos de coral mediante LIBS y las primeras imágenes elementales de alta resolución con una resolución lateral de sólo 15 μm .
- Por último, se exponen las principales conclusiones obtenidas durante los trabajos.

Se han incluido tres anexos:

- En primer lugar diversas consideraciones sobre la seguridad láser en el laboratorio y los efectos de dicha radiación sobre la salud humana.
- En segundo lugar, se mostrará el software e interfaces de usuario generadas para la realización de esta tesis.



Samuel Moncayo Martín

- Por último, se presenta la publicación del trabajo sobre bronce con importancia arqueológica desarrollado durante mi estancia en Bratislava (Eslovaquia).



Abstract

Laser induced breakdown spectroscopy (LIBS) is an analytical technique that relies on the formation and study of a plasma induced by a focused laser pulse on the surface of a sample. The plasma emission produced, due to electronic recombination, generates a spectrum. That spectrum is collected and its spectroscopic analysis allows carrying out qualitative and/or quantitative analysis. LIBS technique has become established as an important spectroscopic technique with enormous analytical potential due to its particular characteristics and advantages over other analytical techniques. Its speed, versatility and the capacity to produce analytical information for almost any type of sample in a minimally destructive way with a reduced or none sample preparation are of utmost importance. LIBS technique is a robust and compact methodology with an easy implementation that allows to perform analytical measurement in situ and even remote analysis.

The main objective of this thesis is the study of the LIBS technique in combination with chemometric analysis methods, especially artificial neural networks, in order to carry out several qualitative studies using LIBS spectra as characteristics spectral fingerprints of the samples. LIBS technique is widely known as well as chemometric methods; however its combination is nowadays under development and only several authors use advanced chemometric methods to improve the performance and the possibilities of the technique. Herein, an important contribution to the study and implementation of several chemometric methods for the analysis has been made, creating new applications of the LIBS technique. LIBS has been applied to analyze samples with interest in the food industry, forensic science and environmental science.



The results obtained provide successful results with a high degree of reliability and robustness.

To achieved this objective, the thesis has been divided into two parts:

- The objective of the first part (Cap. 1 - 3) is to provide a wide view of the technique as well the chemometric methods from the theoretical and practical point of view lead to the analysis and classification of different kind of samples.
 - Specifically, in chapter 1 “*General Introduction*” a brief comment of the fundamental aspects of the laser radiation, plasma emission and the characteristics and theoretical considerations of LIBS have been discussed. A summary of the most important LIBS application has been also included.
 - The analytic aspects of the LIBS technique, in terms of both qualitative and quantitative analysis were discussed in chapter 2 “*Analytical LIBS aspects: qualitative and quantitative analysis*”. Different types of qualitative analysis and methods are discussed, including the calibration based on physical plasma parameters (CF-LIBS). The figures of merit of both quantitative and qualitative analysis have also been considered in this chapter.
 - All the classification chemometric methods used in this thesis are discussed in detail from a theoretical point of view in chapter 3 “*Chemometrics methods applied to the analysis LIBS*”. Linear unsupervised classification methods such as principal component analysis (PCA), and linear supervised classification methods such as discriminant analysis (LDA), Soft Independent Modeling of Class



Analogies (SIMCA), Partial Least Squares-Discriminant Analysis (PLS-DA), Classification and Regression Tree (CART), Binary Logistic Regression (BLR) have been studied. Moreover, nonlinear classification methods such as artificial neural networks (NN) and Support Vector Machines (SVM) have been also included. Due to the importance of the artificial neural networks in the development of the thesis, in this chapter a final section was dedicated to its study, the most important architecture, functions and algorithms have been considered. This section gives a different vision of the neural networks, treating in detail its operations.

- In the second part (Chapter 4 - 7), the most important experimental results in the development and implementation of the LIBS technique to the classification of several samples are shown.
 - Chapter 4 "Instrumentation and experimental results" details all the experimental set up used and the LIBS instrumentation.
 - Chapter 5, "LIBS and chemometric methods applied to the food analysis" deals with the development of a new methodology based on LIBS and chemometric methods for detecting adulteration in food products, in particular olive oils in order to ensure the quality and control of the denomination of origin of Spanish red wines.
 - In Chapter 6, "LIBS applied to forensic science" have been studied and the performance of seven chemometric methods in the classification of samples with interest in the forensic science by means of LIBS spectra has been compared. The combination of LIBS with several chemometric



methods has attracted a huge interest in the LIBS community and in this chapter we propose a system to evaluate different chemometric methods in the classification of samples based on LIBS measurements.

- Finally, chapter 7 "LIBS applied to environmental analysis" show the results of the studies carried out using samples of environmental and archaeological interest. The Mg / Ca ratio in mollusk shells by LIBS has been measured for the first time, and the results have been compared with stable isotopes of oxygen analysis, looking for the correlation with the surface temperature of the sea. Furthermore, the preliminary experimental results of the environmental study of coral skeletons using LIBS will be also presented together with the very first elemental imaging of coral.

To conclude the most relevant conclusions are presented.

Two annexes have been also included:

- First, a consideration of the laser laboratory safety and the effects of this radiation on human health.
- Secondly, the software and user interfaces generated for the realization of this thesis are shown.
- Finally, a third one showing the publication developed during my research stay in Bratislava (Slovakia) on the analysis of different bronze alloys with special interest in the field of archaeological investigations.



Objetivos

La espectroscopia de ablación laser es una de las pocas técnicas que, por sus características, permite el análisis rápido de muestras en cualquier estado de la materia, produciendo una gran cantidad de información espectral, haciéndola particularmente adecuada para el tratamiento de datos mediante métodos quimiométricos de análisis. Esto justifica el esfuerzo que supone su estudio, y del que pueden plantearse los siguientes objetivos:

1. Estudiar, analizar y comparar métodos quimiométricos que mejoren las capacidades analíticas de la técnica LIBS desarrollando nuevas aplicaciones.
2. Demostrar la capacidad de la combinación entre LIBS y métodos quimiométricos para llevar a cabo análisis cualitativos de identificación y clasificación e investigar el potencial analítico de la técnica en aplicaciones concretas, facilitando las tareas de detección de adulteraciones en alimentos y la reconstrucción de individuos a través de restos humanos en el ámbito forense.
3. Desarrollar una metodología de análisis rápida, sencilla y robusta basada en redes neuronales para la clasificación de muestras en condiciones reales y de procedencia tanto conocida como desconocida.
4. Demostrar la viabilidad de la técnica LIBS en el estudio de la variabilidad climática en el pasado mediante el análisis de moluscos y corales y comparar los resultados con técnicas y métodos ya establecidos.
5. Demostrar el potencial de la técnica para la obtención de imágenes elementales y sus posibles aplicaciones al análisis geológico y paleoclimático en combinación con métodos quimiométricos.



Publicaciones

El desarrollo de esta tesis ha generado la publicaciones de los siguientes trabajos:

Publicación en revistas:

1. Plume dynamics of laser-produced swine muscle tissue plasma. J.J Camacho, L. Díaz, A. Marin-Roldan, S. Moncayo and J. O. Caceres et al. Applied Spectroscopy. In press
2. “Evaluation of supervised chemometric methods for sample classification by Laser Induced Breakdown Spectroscopy” S. Moncayo, S. Manzoor, F. Navarro-Villoslada, J. O. Caceres. Chemometric and Intelligent Laboratory Systems 146 (2015) 354-364.
3. Mg/Ca ratios measured by Laser Induced Breakdown Spectroscopy (LIBS): a new approach to decipher environmental conditions. A. García-Escárzaga, S. Moncayo, I. Gutiérrez-Zugasti, M. R. Gonzalez-Morales, J. Martín-Chivelet, J. O. Caceres. Journal of Analytical Atomic Spectrometry. 30 (2015) 1913-1919.
4. “Discrimination of Human bodies from Bones and Teeth Remains by Laser Induced Breakdown Spectroscopy and Neural Networks”. S. Moncayo, S. Manzoor, T. Ugidos, F. Navarro-Villoslada, J. O. Caceres. Spectrochimica Acta Part B Atomic Spectroscopy 101 (2014) 21-25
5. “Fast Bacterial Identification by Laser Induced Breakdown Spectroscopy”. Industrial, Medical and Environmental Applications of Microorganisms: Current Status and Trends.” S. Manzoor, S. Moncayo, F. Navarro-Villoslada, J. A. Ayala, R. Izquierdo-Hornillos, F. J. Manuel de Villena, J. O. Caceres. Wageningen Academic Publishers (2014)
6. “Rapid identification and discrimination of bacterial strains by laser induced breakdown spectroscopy and neural networks”. S. Manzoor, S. Moncayo, F. Navarro-Villoslada, J. A. Ayala, R. Izquierdo-Hornillos, F. J. Manuel de Villena, J. O. Caceres. Talanta 121 65-70 (2014)



7. “Implicaciones de los isótopos estables de oxígeno ($\delta^{18}\text{O}$) y de la relación elemental Mg/Ca en la determinación de la estación de captura de los recursos malacológicos durante el Mesolítico en el Cantábrico”. A. García-Escárzaga, I. Gutiérrez-Zugasti, S. Moncayo, J. Martín-Chivelet, F. J. Manuel de Villena, J. O. Caceres and M. R. Gonzalez-Morales. Actas de congreso Ibérico Arqueometría. 279-288 (2014)
8. “Application of Laser-Induced Breakdown Spectroscopy (LIBS) and Neural Networks to Olive Oils Analysis”. Jorge O. Caceres, S. Moncayo, J. D. Rosales, F. C. Alvira, G. M. Bilmes .Applied Spectroscopy 67(9) 1064-72 (2013)
9. “Determination of the postmortem interval by Laser Induced Breakdown Spectroscopy using swine skeletal muscles”. A. Marín-Roldan, S. Manzoor, S. Moncayo, F. Navarro-Villoslada, R.C. Izquierdo-Hornillos, J.O. Caceres. Spectrochimica Acta Part B Atomic Spectroscopy. 88 186-191 (2013).

Patente:

1. P201400880 titulada “Método de análisis de bebidas alcohólicas”.

Contribuciones en congresos internacionales:

1. 14 - 18 Septiembre, 2015. Euro-Mediterranean Symposium on LIBS. Linz (Austria)
“Evaluation of supervised chemometric methods for sample classification by Laser Induced Breakdown Spectroscopy”. (Póster)
2. 3 - 5 Junio, 2014 Charles University in Prague. WDS 2014
“Discrimination of copper alloys with archaeological interest using LIBS and Chemometric Methods”. (Póster)



3. 19 -21 Junio, 2013. 2º Congreso Internacional de Seguridad Alimentaria. Madrid
“Aplicación de Espectroscopia de Ablación Láser (LIBS) y Redes Neuronales (NN) al análisis y control de adulteración en Aceites de Oliva” (Póster).
4. 11 - 15 Septiembre, 2011. Euro-Mediterranean Symposium on LIBS. Izmir (Turquía)
“Identification and Discrimination of Bacterial Strains by Laser Induced Breakdown Spectroscopy and Neural Network”. (Presentación Oral).
5. 22 - 24 Junio, 2011. Healthcare Systems Ergonomic and Patient Safety (HEPS). Oviedo
“Discriminación de cepas bacterianas mediante LIBS y NN” (Póster).

Contribuciones en congresos nacionales:

1. 2 - 4 Septiembre, 2015. III Simposio de Jóvenes Investigadores en Espectroscopia Aplicada. “evaluación de métodos quimiométricos supervisados para la clasificación de restos óseos humanos mediante espectroscopia de ablación láser (LIBS)” (Presentación Oral).
2. 27 – 28 Febrero, 2014. Conferencias LIBS Inviernos 2014. Madrid. “Métodos y estrategias para llevar a cabo una determinación cuantitativa utilizando espectroscopia láser (LIBS)” (Presentación Oral).
3. 15 – 19 Julio 2013, Curso de Verano Complutense, El Escorial. “Neural Networks as a tool to support the investigation data analysis” (Presentación Oral).
4. 17 -20 de Septiembre, XXIII Reunión Nacional de Espectroscopia, Córdoba. “¿Cómo puede la quimiometría mejorar un análisis LIBS? Aplicación de métodos quimiométricos en la discriminación de individuos a través de restos óseos y dentales” (Póster).



Capítulo 1

Introducción General

El estudio de la materia y su caracterización ha sido objeto de estudio de la física desde sus inicios. Actualmente se conocen muchos estados de la materia además de los cotidianamente conocidos “sólidos, líquido y gaseoso”.¹ Algunos de estos estados son generados en laboratorios o están presentes bajo circunstancias particulares. Entre ellos se encuentra el plasma que, a pesar de no ser comúnmente conocido por la sociedad, es el estado más abundante del universo, estando presente en el sol, estrellas, nebulosas, etc. Sin embargo en la Tierra es mucho más escaso y por tanto menos conocido. El plasma se puede definir como un medio parcial o totalmente ionizado, en el que coexisten electrones libres, aniones, cationes y átomos neutros. La densidad de cargas positivas y negativas es similar y por tanto se puede considerar electrónicamente neutro. El plasma fue descrito por primera vez en los años 20 por Irving Langmuir (premio nobel en Química en 1932). Sin embargo, fueron otros científicos los que comenzaron el estudio de plasma, en particular Benjamin Franklin en 1752, Michael Faraday en 1820, William Crookes 1880, quien definió por primera vez la observación de un gas radiante, J. J. Thomson en 1897 con el descubrimiento de los rayos catódicos. El comportamiento del plasma es similar al de los gases, con la excepción de que dentro del plasma las partículas tienen un comportamiento colectivo, encentrándose ligadas a través de fuerzas electromagnéticas generadas entre ellas, siendo un buen conductor eléctrico¹. Otra importante diferencia es la posibilidad de confinamiento lejos de las paredes, así como su manejo espacial. Dentro de un plasma se desarrollan varios procesos, entre ellos cabe destacar tres:²



- En primer lugar “la ionización y/o excitación”, debido a la alta densidad de electrones libres, estos pueden interactuar con átomos neutros generando especies ionizadas/excitadas.
- Dicha ionización, produce el segundo efecto, la ionización en cadena, produciendo en algunos casos una corriente eléctrica.
- En tercer lugar se produce la emisión de luz por parte del plasma, debido al decaimiento energético y cuyo análisis espectroscópico permite conocer las especies contenidas en el plasma. Además de todos estos procesos se pueden producir disociaciones moleculares, reactividad con superficies, etc.

Las magnitudes más importantes de un plasma son tanto su densidad electrónica como su temperatura. Estas magnitudes permiten definir diferentes plasmas, ejemplificados en la [Fig. 1.1](#), los cuales generan una multitud de aplicaciones como la iluminación por plasma (lámparas de arco), esterilización de materiales que no soportan altas temperaturas, funcionalización de superficies, reactores de fusión nuclear por confinamiento (Tokamak) y por su puesto estudios espectroscópicos.

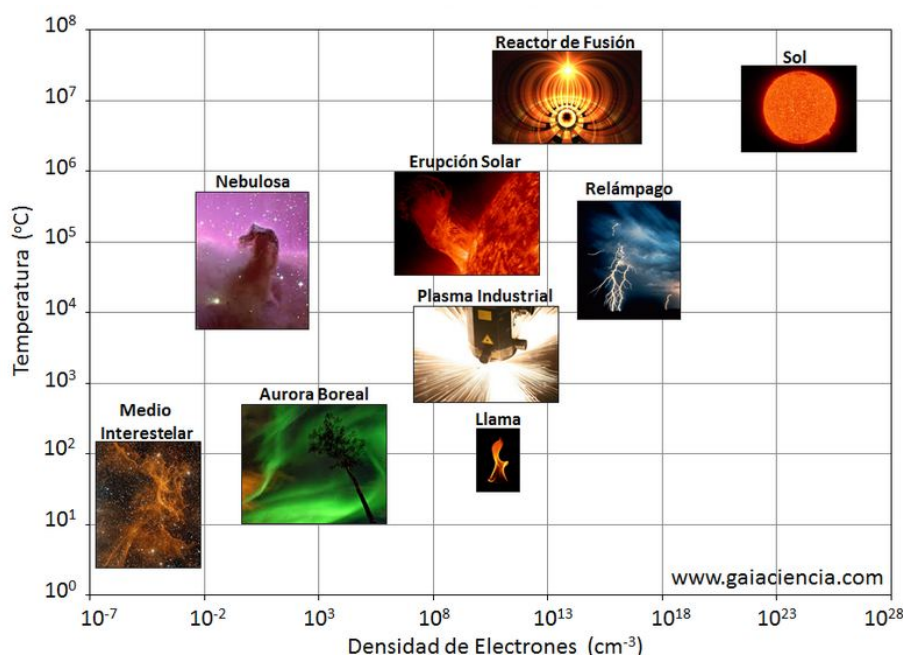


Figura 1.1: Clasificación de plasmas en función de su densidad y temperatura electrónica.

1.1 Espectroscopia de ablación láser (LIBS)

Desde su invención en el año 1960^{3,4} el láser ha sido “una solución en busca de problemas”, siendo la invención más versátil del siglo XX, estando presente en la resolución de problemáticas en un gran ámbito de aplicaciones, especialmente cuando es necesario el uso de fuente de luz controlada y localizada.

Una importante aplicación de la tecnología láser, es la formación de plasmas inducidos. En este caso la exposición de un material a un haz láser focalizado, produce su ablación y la formación de un plasma de alta temperatura y densidad electrónica. La espectroscopia de ablación láser (LIBS) hace uso de la emisión óptica de dicho plasma para obtener información característica de las especies presentes y resolver problemas de tipo cualitativo y cuantitativo. La técnica LIBS se conoce desde la década de los 60, el primer trabajo publicado que usaba un láser para la producción de un plasma data del año 1962 por Brech and Cross.⁵ La obtención del plasma sólo era posible en aire y no en gases inertes, debido al uso de láseres continuos sin la suficiente energía. Las



aplicaciones analíticas de la técnica en ese momento eran pocas ya que la pluma del plasma era demasiado débil para la obtención de un espectro con utilidad analítica. Se realizaron varias mejoras con el fin de aumentar el poder de excitación lo que lleva asociado un aumento en el tamaño y la luminosidad de la pluma generada. Para ello se introdujeron al sistema experimental una serie de electrodos, el primer trabajo donde se describe dicha metodología se debe a John Maxwell en 1963.⁶ En ese momento, además se introdujeron mejoras en el sistema experimental, como el uso de un láser de rubí dotado con sistema conmutador (*Q-switch*, en inglés) para la obtención de pulsos cortos y por primera vez originó un espectro analíticamente útil para el análisis de muestras de interés geológico. Rápidamente se observó el potencial de la técnica para el análisis multielemental y el número de trabajos creció rápidamente entre los años 1964 y 1965. En 1964 R. C. Rosan et al.⁷ presentaron el primer espectro LIBS obtenido sin hacer uso de electrodos auxiliares. Haciendo uso de la técnica sin electrodos, Runge⁸ determinó Ni y Cr en acero inoxidable usando un láser de rubí pulsado, siendo esta la primera publicación LIBS tal y como conocemos la técnica hoy en día. En dicho trabajo se incluyó la primera curva de calibración LIBS, demostrando así su potencial en análisis de tipo cuantitativo.

La técnica llegó a Europa y en 1963, Debras-Guédon y Liodec⁹ publicaron el primer estudio espectroscópico de un plasma generado por ablación láser pulsada (100µs de pulso, 100 J de energía capaz de producir spot de 100 µm de tamaño). Además de líneas atómicas, los autores observaron bandas de emisión de especies moleculares como CN y AlO. Las aplicaciones geológicas fueron sin duda las que más se desarrollaron, la técnica permitía un análisis multielemental semicuantitativo, lo cual era difícil de obtener a través de otras técnicas, en aquel momento, y la técnica rápidamente fue



exportada a otros campos como la minería, arqueología, biología, ciencias forenses y biología entre otras.

En 1981, T. R. Loree et al. con su trabajo "*Laser-induced breakdown spectroscopy: Time-integrated applications*"¹⁰ rebautizaron la técnica con el acrónimo LIBS y a partir de ese momento empezó a cobrar una mayor importancia para los investigadores, siendo muchos los que comenzaron a trabajar en ella permitiendo un gran avance tanto en su conocimiento como sus aplicaciones. En la actualidad es una técnica en auge que sigue despertando un enorme interés gracias a las mejoras instrumentales y la tecnología actual, que permiten el uso de láseres, espectrómetros y detectores más económicos y compactos.

1.2 Características de la técnica LIBS

La espectroscopia de ablación láser es una técnica sencilla y rápida que permite realizar un análisis químico cualitativo o cuantitativo multielemental, con una sensibilidad razonable (del orden de ppm para todos los elementos) a muestras en cualquier estado de agregación, con una baja o nula preparación de muestra. La obtención del espectro es instantánea y por tanto puede ser utilizada para el análisis de muestras en tiempo real. Permite el análisis de materiales tanto conductores como no conductores in-situ, sin limitación de tamaño de la muestra. Es una técnica altamente dinámica, la cual permite el análisis de elementos mayoritarios y en concentraciones traza con un único pulso láser, con una alta resolución lateral y en profundidad (5 – 100 μm) con la capacidad de realizar análisis mínimamente destructivos (ng de material con cada pulso).

Otra ventaja adicional de la técnica es su sencilla implementación para colaborar con otras técnicas, como *Laser Induced Fluoresce* (LIF), *X-Ray Fluoresce* (XRF), Raman ó



Spectral Imaging, entre otras. La [Fig. 1.2](#) muestra el comportamiento de la técnica LIBS frente a los requerimientos ideales de las técnicas espectroscópicas. Sin embargo a día de hoy, la técnica presenta algunos problemas intrínsecos que han impedido una mayor expansión en la industria y en análisis rutinarios como: Análisis micro destructivo, la baja reproducibilidad (necesaria experiencia), sensibilidad limitada (ppm para la mayoría de los casos), baja correlación entre diferentes laboratorios y una alta influencia de los elementos presentes en la matriz. Como consecuencia no existen procedimientos estándar de análisis y cada aplicación necesita de un desarrollo y optimización previa.

A pesar de sus limitaciones intrínsecas, existen formas de mejorar y superar la mayoría de las inconvenientes, por ejemplo, la aplicación de sistemas doble pulso mejoran la sensibilidad y límites de detección de la técnica. Las ventajas operacionales que presenta sobre el resto de técnicas para realizar análisis in-situ a alta velocidad, hacen que LIBS sea y siga siendo una técnica analítica importante con un gran potencial por descubrir y con multitud de aplicaciones por desarrollar.



Figura 1.2: Comportamiento de la técnica LIBS frente a los requerimientos ideales para una técnica espectroscópica

1.3 Fundamentos de la técnica LIBS

El proceso LIBS consiste básicamente en la irradiación de la superficie de una muestra con una radiación láser pulsada focalizada con energía suficiente (en el orden de los mJ) y con una duración típica de pocos nanosegundos. Después de focalizar el pulso láser la densidad de energía que incide sobre la muestras es del orden de varios GW/cm^2 .¹¹ La interacción entre la radiación láser y la muestra provoca la vaporización del material produciendo un plasma luminoso. El plasma está constituido por una gran densidad de electrones libres, átomos neutros y especies ionizadas. Los procesos de relajación electrónica dentro del plasma producen una emisión de radiación con información física y analítica de la composición de la muestras útil para llevar a cabo análisis espectrales de tipo cualitativo y cuantitativo.¹² Desde el punto de vista cualitativo, la emisión de cada elemento es característica y única y está relacionada con

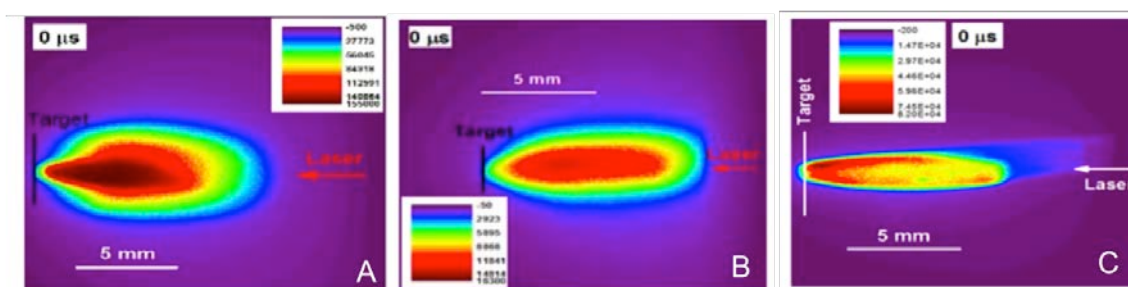


Figura 1.3: Imagen de la pluma a diferentes presiones. A) $7.5 \cdot 10^{-5}$ Torr, B) $2.8 \cdot 10^{-2}$ Torr y C) 750 Torr (presión atmosférica en Madrid).

1.3.1 Proceso de ablación

El término ablación se refiere al proceso a través del cual una pequeña cantidad de materia es evaporada de la superficie de un material, o en alguna excepción del seno de un líquido, mediante la irradiación con una fuente de láser de suficiente energía, produciendo la generación de una emisión óptica y sonora.²⁰

Cuando la energía del láser es suficientemente alta como para romper el umbral de ablación del material, este es rápidamente vaporizado, los enlaces químicos se rompen produciendo átomos e iones en estados excitados generando un plasma de alta temperatura, en torno a los 15000 K.²¹ Este cambio de temperatura ocurre a alta velocidad, 1000 Km/s, lo cual provoca una diferencia de presión de unos 10^8 Pa entre el plasma y los alrededores, produciendo una onda de choque sonora apreciable por el usuario. En algunos casos, dicha onda de choque puede ser registrada y usada analíticamente, lo cual se conoce como “*photoacoustic induced by laser ablation*”.^{22,23}

El umbral de ablación se define como la potencia mínima por unidad de área necesaria para producir la ablación de un material.¹² Por lo tanto el proceso de ablación no solo depende de las características y propiedades propias del material a estudiar sino también de las características del láser usado, como son su longitud de onda y de la duración del



pulso. En cuanto al material, el umbral de ablación depende de su absorbancia a la longitud de onda de trabajo, la fortaleza del enlace y su tensión de vaporización. A modo de ejemplo, el umbral para un metal está en torno a 0.5 GW/cm^2 , en el caso de rocas, cerámicas el umbral está próximo a 1 GW/cm^2 y en el caso de materiales orgánicos en torno a 2 GW/cm^2 .

Se han propuesto dos posibles mecanismos para explicar el fenómeno de ablación (con pulsos ultracortos y con pulsos cortos) la principal diferencia entre ambos se debe a la duración del pulso láser incidente.¹⁹ Es importante destacar que en el plasma las colisiones entre electrones y fotones así como los procesos térmicos están en el orden de los picosegundos, y por tanto en función de la duración del pulso láser, su interacción con el material y el plasma inducido es diferente y el proceso de ablación se produce de forma distinta. Los pulsos ultracortos son pulsos láser de duración menor a los picosegundos, generalmente se lleva a cabo con láseres de femtosegundo que trabajan en un dominio del tiempo menor a los tiempos de difusión del calor. Por otra parte, pulsos cortos, de duración superior a picosegundos, normalmente del orden de pocos nanosegundos, su duración hace que el haz láser interaccione con el material y el plasma inducido produciendo efectos térmicos. Este tipo de láseres son los más habituales y utilizados en LIBS debido tanto a su menor coste como a su mayor facilidad de uso.

1.3.2 Ablación con pulsos cortos

Se trata de pulsos cuya duración supera los tiempos de termalización. La energía electromagnética es transferida a los electrones instantáneamente, y por tanto siguen ganando energía durante toda la duración del pulso transfiriéndola a los alrededores, produciendo cráteres sin bordes definidos y donde es posible apreciar material

solidificado sobre los bordes y en el fondo del cráter, que proviene de la ablación. En la [Fig 1.4](#) se puede observar un cráter producido por un láser de Nd:YAG con un pulso de duración de 5 ns. El principal mecanismo, que rige este proceso, es la llamada ablación térmica o sublimación directa. Primero, se produce la fusión del material y después una explosión del material donde los electrones e iones están en equilibrio térmico. Esto implica que el material extraído no solo proviene de la zona irradiada sino también de los alrededores afectados por el calor.

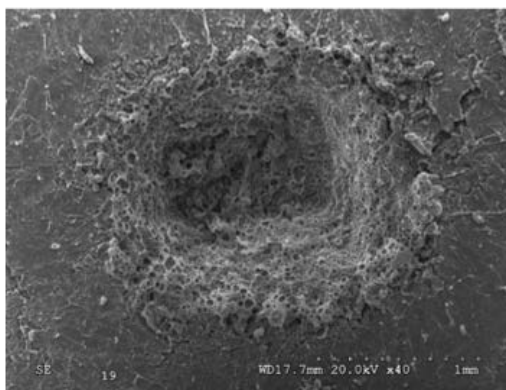


Figura 1.4: *Imagen SEM del cráter producido por un láser de Nd:YAG en una muestra de tejido de musculo porcino.*

1.3.3 Ablación con pulsos ultracortos

Son pulsos producidos por láseres de femtosegundo, donde los tiempos de las colisiones entre electrones y fotones y la difusión del calor son varios órdenes de magnitud mayores que la duración del pulso. En este caso, los cráteres presentan una geometría cilíndrica con bordes bien definidos y prácticamente sin material solidificado en las proximidades del plasma. El material ablacionado solo proviene de la zona irradiada, no se ve afectado por las zonas colindantes y la cantidad ablacionada sólo depende de la profundidad de penetración del haz y de su diámetro. La [Fig. 1.5](#) muestra

un cráter generado por interacción de un pulso femtosegundo de Ti:Zafiro a 430 mW, 800 nm, y con pulsos de 35 fs sobre la superficie de un metal.

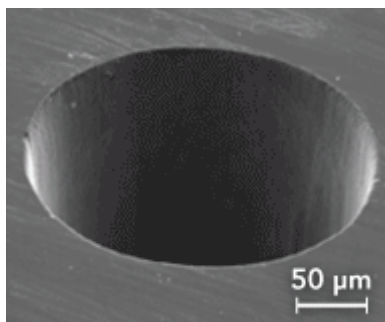
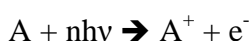


Figura 1.5: Imagen SEM del cráter producido por un láser Ti:Zafiro sobre la superficie de un metal. (Handbook Cremers et al.¹²)

1.3.4 Mecanismos de ablación

En ambos casos, pulsos cortos y ultracortos, los principales mecanismos que sustentan el plasma generado son la ionización/recombinación y la excitación por impacto electrónico. La ablación o *breakdown* es un proceso complejo que se inicia por una ionización multifotónica y se propaga por una ionización en cascada. Este proceso se puede resumir con el siguiente esquema:²⁰



Donde A y A^+ son un átomo y un ion respectivamente que provienen del material y $nh\nu$ es la energía de n fotones. Cuando la absorción de varios fotones, supera la energía de ionización del material, en un proceso llamado ionización multifotónica (MPI), se producen electrones libres que producen la sucesiva ionización de átomos del material. La generación de los primeros electrones libres (*seed electrons*) pueden ser debidos a posibles impurezas de especies con baja energía de ionización en el punto focal del láser.²⁰ Las colisiones entre los electrones libres y los átomos neutros, son responsables



de la producción de una cascada de electrones y la generación de un plasma. Posteriormente, otro fenómeno a tener en cuenta es la explosión de Coulomb,²⁴ la cual consiste en la expulsión de iones después que los electrones han sido expulsados del material mediante un proceso MPI. Este fenómeno ocurre cuando los electrones expulsados no son remplazados instantáneamente (orden de picosegundo) por los electrones próximos, creándose una diferencia de carga superficial en la que los iones sufren una gran repulsión. Si la repulsión electrostática sobrepasa la energía de enlace, los iones son expulsados del material restableciéndose la neutralidad. Este proceso no es frecuente cuando se irradian materiales conductores, ya que la neutralidad se establece rápidamente, sin embargo, en el caso de materiales dieléctricos y semiconductores se produce con mayor frecuencia.²⁵ El proceso de ablación es complejo y no se puede realizar una separación clara entre un tipo de mecanismo u otro, sino que ambos coexisten. Tanto el mecanismo de ablación térmica y la cascada de ionización, como la explosión de Coulomb están altamente influenciados por el tipo de material bajo análisis, así como de las características del láser de excitación (longitud de onda, duración del pulso). Se puede concluir que, la ablación térmica se produce cuando se irradia con pulsos largos y longitudes de onda cercanas al IR en el análisis de materiales conductores, mientras que la ablación multifotónica y la explosión de Coulomb, se producen con longitudes de onda cortas, pulsos cortos e irradiancias altas en materiales dieléctricos.

Un aspecto importante cuando se utilizan pulso cortos, es el fenómeno que ocurre justo después de la ablación del material. La energía del láser no es transferida directamente a la muestra, los fotones incidentes interaccionan con los electrones libres produciendo su calentamiento y provocando el efecto conocido como *Bremsstrahlung* inverso. El efecto



Bremsstrahlung es de vital importancia en los primeros instantes del plasma, produce un continuo de radiación que impide la observación de líneas atómicas e iónicas del plasma, y por tanto es necesario generar un retardo entre la formación del plasma y la adquisición del espectro. Otro efecto a tener en cuenta es la reflexión parcial del haz laser por parte del plasma, este efecto escudo es también conocido como apantallamiento o “*shielding*”. Una vez que se ha formado un gas ionizado, este se expande hacia los alrededores calentando las capas de gas adyacentes haciéndose opaco a la radiación incidente e impidiendo la llegada de la radiación a la superficie de la muestra. Este efecto se ve favorecido con el uso de láser en el IR frente a los láser de UV.¹⁸

Hasta este punto, todos los mecanismos han sido descritos para muestras sólidas, sin embargo, una de las principales ventajas de la técnica LIBS, es la posibilidad de ser aplicado a muestras en cualquier estado de agregación. En el siguiente párrafo se describe el proceso de ablación para muestras en estados de agregación diferentes a la fase sólida, como son los líquidos, gases y aerosoles de gases.

Los estudios LIBS aplicados a muestras líquidas han sido ampliamente investigados para el desarrollo de aplicaciones principalmente médicas. Cuando se trabaja con muestras biológicas, el líquido fundamental que constituye la mayor parte de los tejidos es el agua.

Generalmente el umbral de ablación de líquidos es superior al encontrado en sólidos, esto es debido principalmente a una menor sección de ablación que dificulta la formación del plasma a través de mecanismos multifotónicos.¹¹ Este efecto se traslada al caso de gases, donde los umbrales de ablación son aún más altos. En gases puros, la cantidad de electrones libres (semilla) en la zona focal del láser es muy baja, por tanto la

producción del plasma está poco favorecida.²⁶ Al igual que ocurría en sólidos, la presencia de impurezas bien sean líquidos o partículas en el gas (aerosoles) favorecen el inicio de la ablación. Se pueden producir diferentes zonas susceptibles de inicio del plasma que finalmente colapsan en un único plasma. El efecto de autofocalización que se produce en los líquidos puede favorecer la formación del plasma, haciendo el haz láser más delgado, generando una mayor irradiancia y favoreciendo la ablación.

1.4 Evolución del plasma inducido por láser

Una vez el plasma ha sido creado por cualquiera de los mecanismos mencionados anteriormente, este, evoluciona durante un cierto tiempo debido a las colisiones y procesos radiativos, expandiéndose hasta su completa extinción. Este proceso de expansión y extinción es función de los propios parámetros y características del plasma como son su temperatura y densidad electrónica, así como las propiedades del medio que lo rodea como la presión y composición del gas ambiente.^{27,28} Debido a la influencia de multitud de parámetros en la dinámica del plasma, su estudio y su evolución se rige por mecanismos complejos, en ocasiones difíciles de entender y no conocidos del todo aún. En este caso y para simplificar el desarrollo, vamos a suponer que el comportamiento del plasma es similar al de su densidad electrónica. De este modo se pueden producir tres comportamientos diferentes:

- $N_e > 10^{19} \text{ cm}^{-3}$, estos altos valores de densidad electrónica se pueden obtener instantes después de haberse producido la ablación y en el inicio de la formación del plasma en condiciones atmosférica o bajo altas presiones. La presión externa impide la expansión del plasma, quedando confinado a una pequeña región del espacio. En este primer momento, el espectro obtenido se encuentra formado mayoritariamente por un continuo de radiación y alguna línea ancha, lo cual

impide en gran medida su estudio espectroscópico no siendo útil para un análisis espectroscópico.

- $10^{16} < N_e < 10^{19} \text{ cm}^{-3}$, estas condiciones se encuentran habitualmente cuando se estudia el plasma a tiempos mayores, del orden de cientos de nanosegundos, después de la ablación. Bajo estas circunstancias, el espectro de emisión obtenido está dominado por las líneas de emisión de las especies presentes en el plasma y se comportan siguiendo la función de distribución de Boltzmann. En ese momento, se considera que el plasma se encuentra bajo equilibrio termodinámico local (LTE) y el espectro es útil para su análisis químico mediante LIBS.
- $N_e < 10^{16} \text{ cm}^{-3}$, no podemos considerar el plasma bajo LTE, hay un desajuste entre las poblaciones de los niveles energéticos bajos frente a los altos. En este momento empiezan a ser importantes las reacciones entre las especies presentes en el plasma con los gases presentes en los alrededores, produciéndose especies moleculares y cuya emisión es visible en el espectro en forma de bandas.

En la [Fig. 1.6](#) se muestra la escala temporal en LIBS junto con la evolución temporal de la banda violeta de la molécula de CN, pudiendo observar la estructura vibracional en un plasma producido a presión atmosférica.

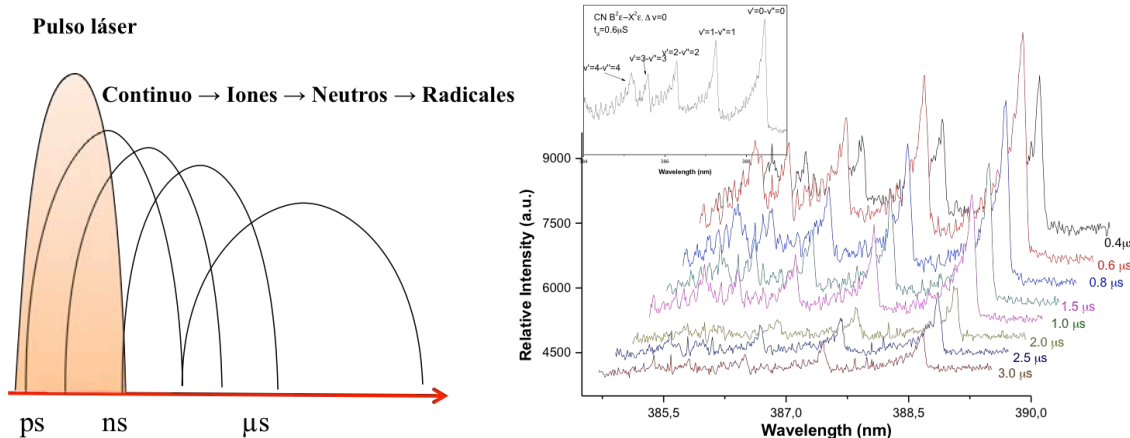


Figura 1.6: Escala temporal y evolución temporal de las bandas vibracionales de la molécula de CN producidas por un láser Nd:YAG a 1064nm en una muestra de tejido porcino.

1.5 Aplicaciones LIBS

La capacidad de producir la ablación de la mayoría de materiales y la formación de un plasma característico de la muestra tras la interacción del pulso láser sobre la superficie del material combinado con las técnicas espectroscópicas de emisión óptica hacen de LIBS una técnica versátil y útil tanto para análisis cualitativo como cuantitativo de una gran variedad de muestras en diferentes disciplinas. Las ventajas más significativas de la técnica LIBS sobre otras técnicas analíticas son: la posibilidad de análisis con escaso o nulo tratamiento de muestra, análisis en tiempo real a alta velocidad, análisis in situ y su carácter mínimamente destructivo. Aunque los procesos físicos que tienen lugar durante la ablación son complejos y hoy en día no están completamente descritos, la versatilidad de la técnica y sus ventajas operacionales han hecho que LIBS, durante los últimos 50 años, haya generado la publicación de miles de artículos, libros y cientos de revisiones sobre los fundamentos teóricos de la técnica, así como las múltiples aplicaciones desarrolladas. En cuanto a las posibles aplicaciones de la técnica LIBS, desde sus inicios las publicaciones usando la técnica LIBS han seguido

un incremento exponencial, la Fig. 1.7 muestra la tendencia desde los años noventa hasta la actualidad. Gracias a su enorme versatilidad las aplicaciones desarrolladas van desde un estudio físico del plasma generado hasta aplicaciones espaciales, pasando por el análisis de metales^{15,29}, líquidos³⁰⁻³³ o sólidos sumergidos³⁴⁻³⁶, aerosoles³⁷⁻⁴⁰, muestras geológicas⁴¹⁻⁴⁴, detección de explosivos⁴⁵⁻⁴⁸ y sustancias peligrosas para la salud como microorganismos^{13,49-51} y virus⁵², aplicaciones forenses⁵³⁻⁵⁶, análisis de plásticos,^{14,57-58} muestras medioambientales⁵⁹⁻⁶¹, con interés arqueológico⁶²⁻⁶⁴ o incluso biomédicas.⁶⁵⁻⁶⁷

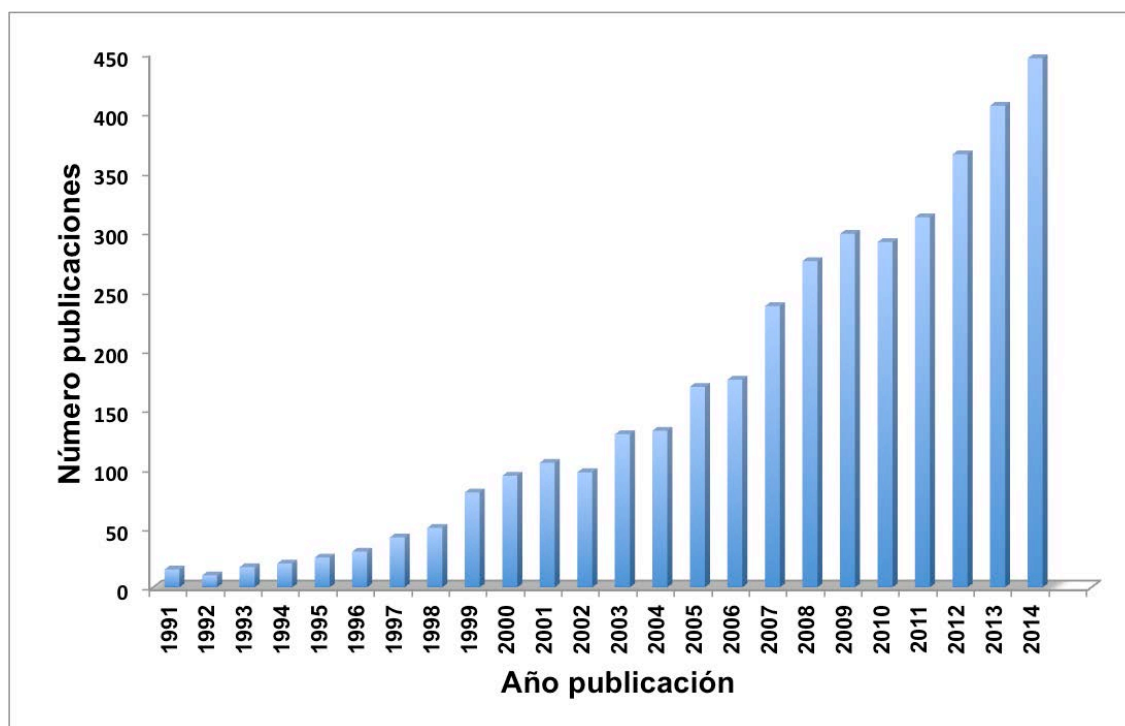


Figura 1.7: Evoluciones de las publicaciones que contienen la palabra LIBS en el desarrollo fundamental de la técnica y sus aplicaciones

Esta figura, no solo indica el incremento del número de publicaciones, también puede indicar el aumento del número de investigadores y por lo tanto de publicaciones en el área, en cualquier caso demuestra que es un área activa y en crecimiento.



Capítulo 2:

Aspectos analíticos de LIBS: Análisis cualitativo y cuantitativo

Una característica importante de la técnica LIBS es la posibilidad de ser aplicada para realizar análisis cualitativos y cuantitativos. Los espectros LIBS pueden ser utilizados como “huella digital espectral” para la clasificación de muestras a través del uso de las herramientas matemáticas adecuada y también es posible realizar análisis de tipo cuantitativo, determinando la concentración de uno o varios elementos simultáneamente, incluso con un único pulso láser. Gracias a su versatilidad, la técnica ha sido desarrollada y aplicada para el análisis de una gran multitud de elementos y matrices.

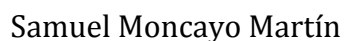
2.1 Análisis cualitativo mediante LIBS

Los análisis de identificación y clasificación de muestras se pueden realizar con diversos objetivos entre los que cabe destacar: la determinación de la ausencia o presencia de un elemento dado, clasificación de una muestra desconocida entre dos a más clases y la evolución de cierto elemento o la estabilidad química de una especie en el seno de una matriz. En el primer caso, para determinar la presencia de un elemento en una matriz, también llamados análisis tipo *screening*, es necesario fijar un límite de concentración por encima del cual podemos asegurar la presencia de dicho elemento con un correspondiente grado de confianza. Este valor límite viene dado fundamentalmente por las características del sistema experimental, en particular por el sistema de detección utilizado, siendo los sistemas con detectores intensificados (ICCD) capaces de obtener menores límites de detección frente a los detectores CCD



convencionales. En general, para llevarlo a cabo se puede usar la intensidad de una única línea o bien de varias líneas debidas a especies tanto atómicas, iónicas como bandas moleculares que estén presentes en el espectro LIBS. Es importante mencionar que cuando la línea o líneas de interés no son detectadas no implica necesariamente que la muestra esté libre de tal elemento, sino que puede encontrarse a una concentración menor al límite de detección de nuestro sistema. Además, es posible que existan interferencias espectrales que pueden producir una identificación errónea. Para este tipo de análisis es deseable contar con un sistema experimental capaz de lograr alta sensibilidad y resolución en la zona de interés analítico.

Para la identificación cualitativa de la composición elemental de una muestra basada en sus líneas de emisión se han aplicado diversos procedimientos, el más comunes es la identificación mediante bases de datos, siendo las más frecuentes las correspondientes a *National Institute of Standard and Technology* (NIST)⁶⁸ y Kuruck (Harvard). Estas bases de datos además de incluir información sobre la longitud de onda de las líneas de emisión incluyen información sobre la intensidad relativa, constantes espectroscópicas tales como los coeficientes de emisión, niveles energéticos del tránsito y degeneración, entre otros. Es habitual que las líneas registradas en las bases de datos no coincidan exactamente con las encontradas en nuestro espectro, por ello es de gran utilidad contar con bases de datos propias. Estas bases propias se pueden obtener con materiales de referencia o bien usando sales inorgánicas de alta pureza y composición perfectamente conocida. Los espectros de referencia han de ser obtenidos en condiciones similares a las de trabajo con el fin de ser representativas. Es especialmente importante mantener una temperatura del plasma similar a la hora de comparar la presencia de especies dentro del mismo.



2.2 Análisis cuantitativo mediante LIBS

LIBS ha evolucionado notablemente en los últimos 40 años gracias, en parte, a la posibilidad de realizar análisis cuantitativos o semicuantitativos multielementales de forma sencilla o incluso in-situ de la mayoría de elementos en diversas aplicaciones, donde se puede destacar los análisis ambientales,⁷⁰ de patrimonio histórico,⁷¹ aplicaciones espaciales⁷² y geológicas,¹⁴ entre otras.⁵⁹ Sin embargo, los problemas asociados a las medidas cuantitativas son el talón de Aquiles de la técnica. En general, la dificultad de realizar una medida cuantitativa está asociada a tres aspectos. En primer lugar la robustez y reproducibilidad de las medidas que son función del sistema experimental utilizado. En segundo lugar, el control de la preparación de muestra (presencia de heterogeneidades y efectos de matriz de la muestra) y por último los problemas relacionados con el análisis de los espectros. Uno de los principales problemas, siendo común a la mayoría de técnicas analíticas, es el llamado efecto matriz.⁷³ En LIBS este efecto está acentuado y se manifiesta en situaciones en las que la intensidad de emisión de un elemento no se comporta de forma proporcional a su concentración. La matriz influye enormemente en la formación del plasma, este efecto se denomina ablación selectiva y modifica las características del plasma como son su temperatura, densidad de electrónica, etc, siendo difícil relacionar directamente la intensidad con la concentración, especialmente cuando se trabaja con concentraciones próximas al límite de cuantificación de la técnica.

Para poder aplicar LIBS como método cuantitativo, hemos de estar de seguros de realizar las medidas bajo equilibrio termodinámico local (LTE) y tener un plasma ópticamente fino. La línea de emisión de a I_{ij} de una línea a la frecuencia ν_{ij} está



Además de estas consideraciones, es adecuado analizar el plasma obtenido a fin de asegurar una ablación estequiométrica y la obtención de un plasma homogéneo. La ablación estequiométrica implica que la vaporización de todos los elementos se produce manteniendo las proporciones originales de los elementos que componen la muestra. Es frecuente la generación de una ablación selectiva donde la concentración de las especies en el plasma depende de las propiedades y características de los elementos y por tanto el plasma generado posee una composición diferente a la muestra. A través de la selección adecuada de los parámetros experimentales como la fluencia, longitud de onda del láser, tiempo de duración del pulso, es posible producir una ablación estequiométrica, produciendo un resultado preciso en la cuantificación. La homogeneidad del plasma es más complicada de controlar. El plasma es un sistema dinámico en constante evolución y con interacción con el medio, produciéndose grandes gradientes térmicos. Dentro del plasma se producen diferencias entre las zonas externas del plasma, más frías y en contacto directo con la atmósfera que lo rodea y la zona interna del plasma, más caliente y menos afectada por el entorno. Por tanto, considerando la parte central del plasma es posible asumir que el comportamiento y composición del mismo es más homogéneo y de mayor utilidad desde el punto de vista de la cuantificación.

Para intentar suplir los posibles inconvenientes de la técnica para el análisis cuantitativo, se han propuesto empleo de distintos métodos de cuantificación, como el de estándar interno, y los basados en los parámetros físicos del plasma. A continuación se describen estos métodos.



El análisis cuantitativo puede realizarse de diferentes formas, la más sencilla y directa es analizando la respuesta de una determinada línea de emisión a una concentración dada. Para generar la curva de calibración se han de realizar medidas de la intensidad de un conjunto de patrones de concentración conocida y con una matriz similar a la muestra problema y representar su correspondiente señal, frente a la concentración. Sin embargo, este proceso no es adecuado para un gran número de casos debido al efecto matriz, el cual hace que el elemento de análisis se comporte de forma diferente en los patrones que en las muestras. Para minimizar dicho efecto, se puede usar la normalización con estándar interno, bien sea utilizando la relación entre dos líneas presentes en la muestra,⁵⁴ una línea de la muestra y la señal de un elemento añadido externamente⁷⁸ o incluso entre la línea de interés y el ruido de fondo.⁷⁹

Además de los métodos de cuantificación convencionales para análisis cuantitativo univariante, es cada vez más frecuente el uso de métodos multivariantes para la cuantificación de muestras mediante LIBS, como es el caso del PLS-DA.

Por último, es posible obtener la concentración de un elemento usando información de los parámetros del plasma generado, en lo que se conoce como “*Calibration Free o Free LIBS* ”.⁸⁰ En dicha calibración no es necesario el uso de curvas de calibración convencionales, sino que la concentración puede ser determinada a través de parámetros físicos del plasma, como su temperatura y densidad electrónica, aplicando algoritmos matemáticos que nos permitan su obtención. Este método de calibración presenta una importante ventaja frente a las calibraciones convencionales y es la posibilidad de determinar de la concentración con solo una muestra sin necesidad de medir patrones de concentraciones conocidas. Como contrapartida, se puede destacar

la necesidad de conocer con gran precisión y exactitud los parámetros físicos del plasma, lo cual es difícil para algunas muestras.

2.3.1 Calibración con estándar interno

El procedimiento de calibración con estándar interno puede ser aplicado de dos formas diferentes, en primer lugar es posible usar como referencia la concentración conocida de un elemento presente en la muestra, lo cual implica el conocimiento previo de su concentración y por lo que desde el punto de vista práctico no es muy útil. Sin embargo, es de especial interés en el caso de muestras formadas por capas o láminas en las que la concentración de un determinado elemento varía en función de la capa. Por ello, conociendo la concentración del elemento en una capa, se permite obtener su concentración en el resto de capas.⁸¹ Para llevar a cabo dicho procedimiento, es necesario el uso de la relación entre dos líneas de emisión del elemento y multiplicarlo por la concentración conocida para obtener la concentración desconocida a través de la siguiente ecuación:⁸²

$$C_{x?} = \frac{I_{x?} C_{xS}}{I_{xS}} \quad (\text{Eq. 2.2})$$

donde $C_{x?}$ y C_{xS} son la concentración del elemento X a determinar y la concentración del elemento X de referencia, respectivamente y donde $I_{x?}$ es la intensidad de emisión del elemento X de concentración desconocida y I_{xS} la intensidad del elemento X de referencia.

Es importante mencionar que la relación de intensidades de las líneas ha de realizarse con tránsitos electrónicos similares en energía, es decir próximos en longitud de onda. La selección de dicha línea debe hacerse cuidadosamente, seleccionando aquellas que presenten una alta relación señal/ruido, evitando líneas de resonancia y líneas con posibles interferentes espectrales. Este procedimiento, es usado



habitualmente en otras técnicas analíticas y es una simplificación de una curva de calibrado convencional con un único punto.

Por otra parte, el método de estándar interno puede ser utilizado para minimizar el efecto matriz cuando se busca determinar la concentración de un elemento presente a una baja concentración dentro de una matriz con un elemento mayoritario. Esto implica asumir que la concentración del elemento mayoritario es constante en toda la matriz, cosa que en muchos casos no ocurre, añadiendo un error sistemático en la medida. Se calcula a través del cociente entre la intensidad del elemento traza y la intensidad del estándar interno como se muestra en la siguiente ecuación:

$$I_{normalizada} = \frac{I_y}{I_s} \quad (\text{Eq. 2.3})$$

donde la $I_{normalizada}$ corresponde al cociente entre la intensidad del elemento traza I_y y la intensidad del estándar I_s . En el capítulo 7 esta metodología ha sido aplicada a la determinación de Mg y Sr en muestras de concha de moluscos (*P. Lineatus*) en dos estructuras biogénicas de carbonato de calcio (calcita y aragonito) para su correlación con la temperatura de la superficie del mar (SST). El resultado obtenido por esta metodología no es la concentración absoluta del elemento sino una concentración relativa de dicho elemento frente a otro de referencia. A pesar de no proporcionar un valor absoluto, es muy útil ya que elimina cualquier influencia de parámetros experimentales, especialmente fluctuaciones en la energía del láser, quedando dicho efecto compensado con la relación de intensidades. Además, no es necesario el uso de materiales de referencia o estándares certificados para su aplicación, los cuales en muchos casos o no existen o son difíciles de obtener, también elimina la etapa de calibración del equipo instrumental que está sujeta a una deriva instrumental.

2.3.2 Calibración basada en parámetros físicos del plasma (*Calibration Free*)

Este procedimiento de calibración fue desarrollado por primera vez en el año 1999 por A. Ciucci et al.⁸⁰ Su desarrollo se produjo debido a la necesidad de corregir las desviaciones que produce el efecto matriz en la determinación de la concentración elemental y sobre todo debido a la dificultad de obtener curvas de calibración para muestras complejas, en las cuales es extremadamente difícil la obtención de patrones similares a las muestras.

Este proceso se basa en el conocimiento de los parámetros físicos de los plasmas inducidos por la ablación láser y puede ser aplicada a la cuantificación de sólidos, líquidos y gases. Para poder realizar este tipo de análisis es imprescindible un estudio del plasma para poder asegurar con certeza que el plasma es representativo de la composición real de la muestras, que es ópticamente delgado y estar bajo condiciones de equilibrio termodinámico local (LTE). Para verificar la obtención de un plasma ópticamente delgado se han usado, en bibliografía, diferentes procedimientos como el propuesto por L.J. Radziemski et al.³⁸, midiendo la intensidad de un multiplete debido a N I, midiendo la intensidad de tres líneas de O I⁸³ o el propuesto por J.B. Simeonsson et al.⁸⁴ donde se verifica midiendo la anchura de una misma línea en distintas posiciones axiales del plasma y distintas energías del láser, sin embargo el más sencillo es evaluar la relación de intensidades de un mismo elemento con la relación de los coeficientes de Einstein para esas transiciones.

La determinación precisa de los parámetros físicos del plasma como su temperatura y densidad electrónica, es fundamental para obtener un valor de concentración exacto. Para poder obtener un valor de temperatura preciso, es necesario cumplir la condición de tener un plasma en LTE ya que afecta directamente al valor de

temperatura obtenido. En LTE la temperatura obtenida por las ecuaciones de Boltzmann y Saha es igual a la de la distribución de velocidades de los electrones libres según Maxwell-Boltzmann. Y además se cumple la ley de Kirchhoff de la radiación, la cual establece la relación entre el coeficiente de emisión y el coeficiente de absorción. Para que asegurar el LTE los procesos colisionales de los electrones del plasma han de predominar frente a los radiativos, siendo necesaria una alta densidad electrónica. Según el criterio de McWhirter⁸⁵ existe una densidad electrónica crítica para la cual los procesos de colisiones son diez veces superiores a los radiativos y debe cumplir que:⁸⁶

$$n_e \geq 1.6 \cdot 10^{12} T^{1/2} (\Delta E)^3 \text{ cm}^{-3} \quad (\text{Eq. 2.4})$$

siendo T la temperatura y ΔE la diferencia de energía entre los niveles de la transición. En general la densidad electrónica crítica se encuentra en torno a los $10^{15} - 10^{16} \text{ cm}^{-2}$. Esta densidad es válida cuando se trabaja a presiones entre 1 - 760 Torr, sin embargo cuando se trabaja a bajas presiones este valor tiene que ser evaluado de nuevo. El criterio de McWhirter es necesario pero no exclusivo para lograr el LTE y en consecuencia, para asegurar el LTE se han de aplicar dos criterios adicionales. El primero de ellos trata de comparar las temperaturas de excitación e ionización obtenidas por la ecuación de Boltzmann y las obtenidas por Saha-Boltzmann.³⁸ El otro criterio se basa en la comparación de la distribución de Boltzmann teórica frente a la experimental.

Una vez obtenido un plasma ópticamente delgado y se satisfacen los criterios que aseguran el LTE, para llevar a cabo una cuantificación “free LIBS” es necesario la determinación de la densidad electrónica y la temperatura del plasma.

Para la medida de la densidad electrónica existen tres procedimientos ampliamente utilizados: el uso del ensanchamiento de líneas Stark (STB)²², el desplazamiento Stark (STS)²⁵ y el obtenido a través de la ecuación de Saha-Boltzmann (SBE).⁸⁸ Es frecuente

que para la determinación de la densidad electrónica se usen elementos que provienen de la atmósfera alrededor del plasma como Ar, He, H, N y O.

En la realización de esta tesis los cálculos de densidad electrónica han sido calculados a través del ensanchamiento Stark y por tanto sólo este método se detallará.

El ensanchamiento Stark se produce debido a los choques entre los átomos emisores con electrones e iones, provocando un ensanchamiento de la línea de emisión y un desplazamiento de la longitud de onda de la línea. El ensanchamiento de la línea no sólo se produce por el efecto Stark, el efecto Doppler también está presente, sin embargo, este suele ser menor que el efecto Stark y puede ser despreciado.

En líneas del átomo de hidrógeno o átomos hidrogenoides, donde el efecto Stark es lineal, la densidad electrónica se puede calcular con la siguiente expresión:⁸⁹

$$N_e = C(N_e, T) \Delta\lambda_s^{3/2} \quad (\text{Eq. 2.5})$$

donde $\Delta\lambda_s$ es la anchura máxima a mitad de pico (FWHM) y $C(N_e, T)$ son los parámetros de impacto que se encuentran tabulados y tienen una baja dependencia con la temperatura y la densidad electrónica.⁸⁹

En general debido a que el ensanchamiento Stark y el impacto electrónico son predominantes en las condiciones de medida de plasmas producidos por LIBS, se puede despreciar el factor de corrección iónico y el ensanchamiento Stark puede calcularse como:

$$\Delta\lambda_{FWHM}^{Stark} \approx 2W \left(\frac{N_e}{10^{16}} \right) \quad (\text{Eq. 2.6})$$

donde W es un parámetro tabulado independiente de la densidad electrónica y sólo tiene una dependencia suave con la temperatura electrónica.

Para átomos con más de un electrón, también es posible relacionar la densidad electrónica con la anchura del pico a través del efecto Stark cuadrático que viene dado por la siguiente expresión:⁷⁴

$$w_{total} \approx [1 + 1.75A(1 - 0.75R)]w \frac{N_e}{10^{16}} \quad (\text{Eq. 2.7})$$

donde w es el parámetro de impacto electrónico, A el parámetro de impacto iónico y n_e la densidad electrónica. Los parámetros w y A dependen ligeramente de la temperatura. R es la relación entre la distancia media de los iones y el radio de Debye.

Para medir con exactitud el ensanchamiento Stark y con ello la densidad electrónica, las líneas usadas han de estar lo más aisladas posible de interferentes y sin autoabsorción.

Otra parámetro importante a determinar es la temperatura del plasma, para ello existen diferentes procedimientos entre los que se puede destacar el uso de la ecuación de Boltzmann, la ecuación de Saha-Boltzmann y la determinación a través de tránsitos moleculares (vibracional y rotacional). Existen otros procedimientos que permiten calcular la temperatura en plasmas que no se encuentran en LTE.⁹⁰

Ecuación de Boltzmann

Es posible obtener la temperatura del plasma a través de la intensidad relativa de dos líneas del mismo elemento en el mismo estado de ionización, asumiendo que se cumple la ecuación de distribución de Boltzmann. Para cada línea de emisión, su intensidad es función de ε (emisividad de la línea integrada sobre el perfil de la línea), λ (longitud de onda de la transición), g que es la degeneración del nivel energético superior y A que es la probabilidad de la transición a través de la siguiente ecuación $y \equiv \varepsilon\lambda/(g A)$ y a través de su relación de intensidades es posible obtener⁸⁹:

$$\ln \frac{y_2}{y_1} = -\frac{1}{kT} (E_2 - E_1) \quad (\text{Eq. 2.8})$$

donde k es la constante de Boltzmann, T la temperatura E_2 y E_1 , el nivel de energía superior e inferior del transito, respectivamente la temperatura obtenida a mediante la ecuación de Boltzmann es comúnmente llamada temperatura de excitación, sin embargo bajo condiciones de LTE ha de ser igual a la temperatura cinética de los electrones. Este método permite obtener de forma sencilla una aproximación de temperatura del plasma, sin embargo conduce a considerables errores en el valor de la temperatura donde su incertidumbre su viene dada por:⁹⁰

$$\frac{\Delta T}{T} = \frac{kT}{E_2 - E_1} \frac{\Delta(y_1/y_2)}{y_1/y_2} \quad (\text{Eq. 2.9})$$

Se han realizado diversos estudios donde se ha demostrado que si kt y $E_2 - E_1$ son del mismo orden de magnitud, el error en la determinación de la temperatura es similar al error en la determinación de la intensidad de las líneas de emisión, siendo en muchos casos errores del 50 -70%.⁹¹

Un método para determinar la temperatura del plasma con una mayor exactitud a través la ecuación de Boltzmann es mediante el uso de un mayor número de líneas de emisión.

La representación gráfica entre $y = \ln \frac{I_{\lambda}^{ki}}{g_k A_{ij}}$ frente a $x = E_k$ provee una recta cuyo

ajuste por mínimos cuadrados permite obtener la pendiente de la recta como $m = -\frac{1}{K_b T}$

y la ordenada en el origen $q_s = \ln \frac{C_s F}{U_s(T)}$, donde I_{λ}^{ki} es la intensidad de la línea integrada,

F es un parámetro instrumental, C_s y $U_s(T)$ la concentración y la función de partición de la especie emisora, respectivamente. A partir del valor de la pendiente es posible obtener la temperatura del plasma y a través de la ordenada en el origen es posible obtener la concentración del elemento representado. Además, es posible representar en una misma gráfica varios elementos, calculando así su temperatura en el plasma y la concentración de los mismos.⁹²

Ecuación de Saha-Boltzmann

Para mejorar la exactitud en la determinación de la temperatura obtenida por el método de dos líneas de Boltzmann y la ecuación de Boltzmann, donde se usan líneas de un único estado de ionización del elemento, la ecuación de Saha-Boltzmann incluye el análisis de líneas de un elemento en diferentes estados de ionización. Con ello la diferencia de energía entre estados es superior, debido a la energía de ionización, la cual es mucho mayor a la energía térmica.⁷⁴ La ecuación de Saha-Boltzmann viene expresada a través de la siguiente ecuación relacionando líneas de emisión de diferentes estados de ionización, la temperatura del plasma y la densidad electrónica:⁹⁰

$$\ln \frac{y_2}{y_1} = -\frac{1}{kT} (E_2 - E_1 + E_\infty - \Delta E_\infty) + \ln \left[2 \frac{(2\pi mk)^{3/2} T^{3/2}}{h^3 N_e} \right] \quad (\text{Eq. 2.10})$$

En esta ecuación, para obtener la temperatura del plasma es necesario conocer la densidad electrónica, en LTE la temperatura de ionización es igual a la temperatura electrónica de la ecuación de Boltzmann. Representando en la misma figura ambas especies (neutra e iónica), es posible llegar a una representación similar a la obtenida por la ecuación de Boltzmann, donde el eje de abscisas y ordenada ha de ser modificado para las especies iónicas. La corrección de la ordenada incluye un término que depende de la temperatura como $\ln(T^{3/2})$ y por tanto la temperatura del plasma se obtiene a través de un procesos iterativo hasta que se la pendiente no se ve modificada. Este proceso puede resultar tedioso pero ofrece incertidumbres en la determinación de la temperatura menores, en torno al 5 - 7%.

La [Figura 2.1](#) muestra la gráfica de Boltzmann para diferentes líneas de He I⁹³ (Derecha) y Saha- Boltzmann (Izquierda) para silicio y aluminio en zeolitas.⁹⁴

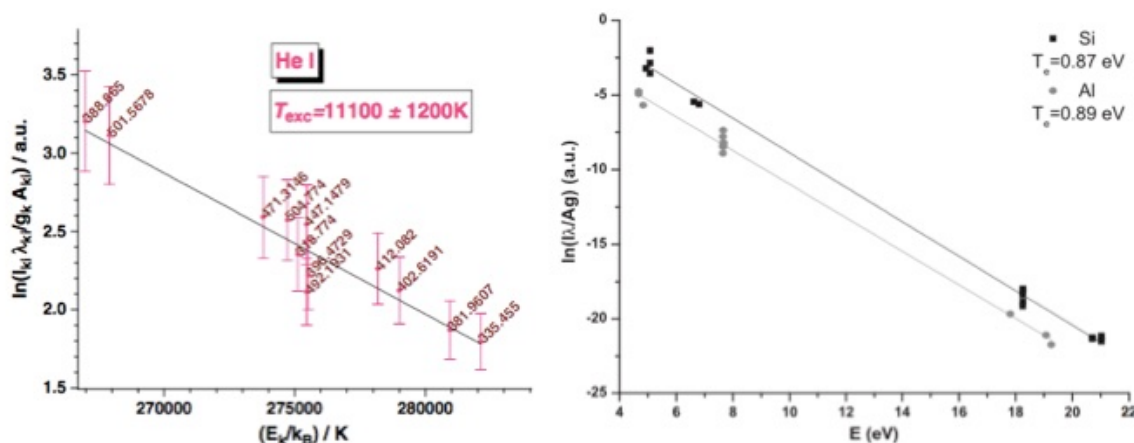


Figura 2.1: Gráfica de Boltzmann para una muestra de gaseosa de He y Saha-Boltzmann para una muestra de zeolita.

2.4 Figuras de mérito de la clasificación

Cuando se realiza un análisis cualitativo es necesario definir que parámetros que han de ser estudiados para evaluar el éxito de la clasificación. Los parámetros más habituales son: la tasa de clasificación correcta, tasa de clasificación incorrecta, tasa de no clasificados, precisión general de la clasificación, sensibilidad, selectividad, capacidad de generalización y por último la robustez.

La tasa de clasificación correcta, se define como el porcentaje de muestras correctamente clasificadas sobre el total, valores altos en este parámetro indican una buena clasificación, sin embargo para obtener un modelo de clasificación apropiado tanto la tasa de clasificación incorrecta, definida como porcentaje entre muestras incorrectamente clasificados frente al total como la tasa de no clasificados, muestras no asignadas a ninguna clase, debe ser baja.



Para definir la sensibilidad y selectividad de un modelo de clasificación es necesario definir la matriz de confusión Tabla 2.1.

Tabla 2.1. Matriz de confusión

		Resultado real	
		Positivo	Negativo
Resultado LIBS	Positivo	Verdadero Positivo (TP)	Falso Positivo (FP)
	Negativo	Falso Negativo (FN)	Verdadero Negativo(TN)

Cuando el resultado de LIBS es positivo y el valor de referencia es también positivo se obtiene un verdadero positivo, simétricamente cuando ambos son negativos se obtiene un verdadero negativo. Por otra parte cuando el resultado LIBS es positivo pero el valor real es negativo se obtiene un falso positivo y de forma similar si el resultado LIBS es negativo pero el real es positivo se obtiene un falso negativo. La sensibilidad (S) y especificidad (SP) se puede calcular como:⁶⁹

$$S = \frac{TP}{TP+FN} \quad (\text{Eq. 2.11})$$

$$SP = \frac{TN}{TN+FP} \quad (\text{Eq. 2.12})$$

un valor de $S = 1$ implica que todas las muestras han sido correctamente clasificadas, la sensibilidad es también llamada tasa de clasificación correcta.

La especificidad, por otro lado es conocida como tasa de clasificación incorrecta. Otro parámetro que se puede calcular a través de la matriz de confusión es la precisión en la clasificación (A) calculado como:

$$A = \frac{TP+TN}{TP+TN+FP+FN} \quad (\text{Eq. 2.13})$$



Un método de clasificación apropiado debe ser capaz de clasificar correctamente todas las muestras incluidas en el entrenamiento (alta sensibilidad), así como clasificar correctamente muestras desconocidas de la misma clase no incluidas en el entrenamiento del modelo (alta capacidad de generalización).⁹⁵

Por último es importante evaluar cómo se comporta el modelo cuando se introduce una muestra ajena al mismo (test de robustez). Idealmente cualquier muestra no perteneciente a ninguna clase incluida en el entrenamiento debe ser clasificada como desconocidas. Para evaluar la robustez del modelo puede realizar un procedimiento conocido como “*leave-one-out*” (LOO) y consiste en eliminar completamente una clase del set de entrenamiento y posteriormente estudiar el comportamiento del modelo frente a dicha clase. Posteriormente cada clase es iterativamente eliminada del entrenamiento y testada. Este procedimiento es considerado una validación del método y será el método y desarrollado con más detalle en el capítulo 6.

2.5 Figuras de mérito de la cuantificación

Las figuras de mérito analíticas son habitualmente usadas para determinar las bondades de un método de análisis y poder comparar su rendimiento con distintas técnicas analíticas a través de cinco parámetros fundamentales: límite de detección y cuantificación, precisión, exactitud, sensibilidad y selectividad.

El límite de detección (L_D) o valor mínimo detectable se define como la menor concentración de sustancia que puede ser detectada con un cierto nivel de confianza.⁹⁶

El límite de cuantificación (L_Q) por su parte, se define como el valor límite de concentración del elemento con un nivel aceptable de exactitud y precisión.⁹⁷ La precisión se define según la guía ISO 2524-1 como el grado de concordancia entre ensayos independientes obtenidos bajo unas condiciones estipuladas y exactitud como

la proximidad en la concordancia entre un resultado y el valor de referencia aceptado. La sensibilidad de un instrumento o un método analítico puede definirse como la capacidad del método para producir un cambio en la respuesta con pequeñas variaciones en la concentración de un analito. Definiéndose matemáticamente como la variación en la señal dividido por la variación en la concentración del analito y por tanto en curvas de calibrado convencionales, la sensibilidad viene definida por la pendiente de la curva.⁹⁸ La selectividad de un método analítico indica el grado de ausencia de interferencias debidas a otras especies contenidas en una matriz. Matemáticamente se define como el coeficiente de selectividad (ξ) del analito con respecto al interferente:⁹⁹

$$\xi_{i,a} = \frac{A_{analito}}{A_{interferente}} \quad (\text{Eq. 2.14})$$

siendo A la pendiente de la recta de calibrado de cada elemento o especie.

2.5.1 Precisión

En LIBS el plasma inducido por láser atomiza la muestra y la excita simultáneamente en cada pulso, por tanto hay una inestabilidad inherente en cada pulso la cual conduce a una pérdida de estabilidad y repetitividad. La medida de la repetitividad se describe en términos de precisión como: la desviación de n medidas (x_i), de la media aritmética M de una serie de repeticiones, esta desviación se calcula como la desviación estándar s por la siguiente expresión:

$$s = \sqrt{\frac{\sum(x_i - M)^2}{n-1}} \quad (\text{Eq. 2.15})$$

y más habitualmente como la desviación estándar relativa (%RSD):⁵⁹

$$\%RSD = 100 \times (s/M) \quad (\text{Eq. 2.16})$$

En general cuando se llevan a cabo medidas LIBS es frecuente realizar varias medidas produciendo un espectro promediado con el objetivo de minimizar la variación pulso a



pulso, debido a diferencias en la excitación del plasma. El número de espectros a promediar ha de calcularse experimentalmente, realizando un gráfico de RSD% frente al número de espectros promediados. El valor óptimo será función de la homogeneidad de la muestra, la precisión con la que queramos llevar a cabo el análisis, del método de análisis y de la relación número de medidas/tiempo. Este procedimiento es frecuente en la mayoría de técnicas analíticas.

2.5.2 Límite de detección y cuantificación

El límite de detección es considerado la cantidad mínima de muestra que puede ser medida con exactitud. Existen diferentes formas de definirlo, la más sencilla es considerar la señal que produce una intensidad igual a la del blanco más tres veces la desviación estándar del blanco. El límite de cuantificación por su parte es considerado como la cantidad de analito que proporciona una señal igual a la del blanco más diez veces la desviación estándar del blanco. El límite de detección se puede ver afectado en cada medida por parámetros como la energía del láser, los tiempos usados para la detección y de las características de la muestra.

2.5.3 Exactitud

La exactitud en LIBS proporciona el grado de concordancia entre el resultado de la medición y el valor de referencia aceptado. El valor de referencia puede provenir de:

- Un material de referencia certificado que asegura que la matriz es adecuada y con una incertidumbre compatible con los requisitos necesarios.
- A través de un método normalizado alternativo, en cuyo caso las muestras no necesariamente han de ser material certificado pero sí debe ser homogénea y su valor ha de ser estimado con su incertidumbre rigurosamente



- A través de un ensayo de intercomparación.

Los errores en la determinación pueden ser sistemáticos cuando hay un fallo en el procedimiento o en el instrumento, estos errores pueden ser detectados y corregidos. Por otra parte los errores aleatorios no se pueden evitar y por tanto forman parte de las propias limitaciones del método. Aunque los errores aleatorios no se pueden eliminar, pueden ser minimizados a través de la realización de réplicas.



Capítulo 3:

Métodos Quimiométricos aplicados al análisis LIBS

3.1 Introducción a la Quimiometría

Los métodos quimiométricos aplicados al análisis y tratamiento de datos se conocen desde principios de los años 70. Sin embargo, los cálculos y algoritmos necesarios para llevar a cabo ciertos análisis quimiométricos han supuesto un hándicap para su desarrollo hasta la aparición de los ordenadores actuales, capaces de procesar y realizar cálculos de grandes conjuntos de datos sin mayor dificultad. Los primeros métodos diseñados específicamente para una aplicación quimiométrica fueron el algoritmo NIPALS (*Nonlinear Iterative Partial Least Squares*), para realizar un análisis de componentes principales (PCA), el método de vecinos próximos y el método SIMCA (*Soft Independent Modeling of Class Analogy*), ambos para reconocimiento de patrones. Actualmente la quimiometría se ha establecido como una metodología madura con un importante papel en el procesamiento de datos generados con los sistemas de medida actuales. Los métodos quimiométricos están extendidos en aplicaciones químicas y más concretamente de la química analítica, sin embargo es aplicada para la resolución de problemas en diferentes ámbitos como por ejemplo el procesamiento de señales, diseño experimental, reducción de variables, exploración de datos, análisis cuantitativo multivariante y reconocimiento de patrones.¹⁰⁰⁻¹⁰²

Uno de los mayores éxitos de los métodos quimiométricos es el análisis de reconocimiento de patrones para la realización de estudios de tipo cualitativo. Existen libros, congresos y revistas especializadas donde se estudia y desarrollan aplicaciones



en el reconocimiento de patrones, abarcando multitud de áreas. Dependiendo del área de conocimiento de los investigadores, el reconocimiento de patrones se ha desarrollado haciendo uso de dos enfoques diferentes. Por una parte, influenciado por la comunidad científica que trabaja con máquinas de aprendizaje (*Machine Learning*) y que incluye el uso de métodos como Redes Neuronales (NN), métodos Kernel, mapas autoorganizados (SOM) y las máquinas de soporte vectorial (SVM). Por otra parte, están los investigadores con un enfoque más estadístico donde se puede destacar los métodos de Análisis Discriminante (DA) y de tipo Bayesiano. Ambos obtienen resultados similares por caminos totalmente diferentes.

En la realización de esta tesis los métodos quimiométricos aplicados van encaminados principalmente al análisis de reconocimiento de patrones usando ambos enfoques. Además se han aplicado métodos para la selección y reducción del número de variables que proporcionen resultados de clasificación satisfactorios, así como un análisis cuantitativo multivariante. La principal diferencia entre los métodos de clasificación y cuantificación es que en los modelos de cuantificación se busca relacionar las muestras a una o más propiedades continuas numéricas, mientras que los modelos de clasificación tratan de buscar la relación matemática existente entre un conjunto de variables descriptivas (medidas) y una variable cualitativa, que define una clase o tipo de muestra.

3.2 Clasificación de los modelos quimiométricos

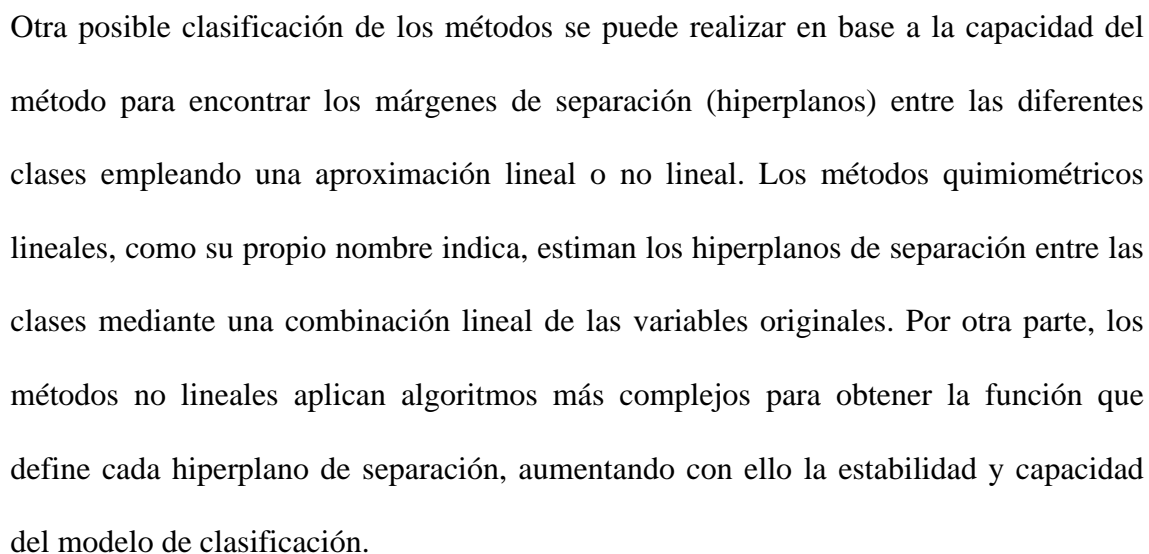
El análisis multivariante es la parte de la estadística y del análisis de datos que estudia, analiza, representa e interpreta los datos que resultan de observar más de una variable estadística sobre un número de objetos. Los métodos de clasificación multivariantes forman parte de una serie de técnicas diseñadas para encontrar modelos

matemáticos capaces de asignar un objeto a su clase de procedencia a partir de en un conjunto de observaciones o medidas.

Existen multitud de métodos quimiométricos, con diferentes características y propiedades disponibles para llevar a cabo un proceso de clasificación.^{103,104} Sin embargo, la selección del método más apropiado en cada caso no es sencilla y los datos deben ser estudiados cuidadosamente con el objetivo de lograr una clasificación satisfactoria.

En cuanto a la clasificación de los métodos quimiométricos que pueden ser usados, no hay un convenio uniforme, sino que surgen diversas clasificaciones en función de las diferentes características de los mismos. Una posible clasificación se puede hacer en base a si los métodos son supervisados o no supervisados. A modo de resumen podemos decir que los métodos no supervisados se definen con una sola matriz de datos \mathbf{X} o matriz de objetos, en la que no se asume ningún modelo de clasificación previo, sino que es el propio método el que genera el modelo, determinando el número de posibles clases y caracterizando las variables que definen cada uno de ellas.

Por otra parte, en los métodos supervisados se definen dos conjuntos de datos, objetos (\mathbf{X}) y clases (\mathbf{Y}). Es importante destacar que constan de dos etapas: el entrenamiento y la validación. Durante el entrenamiento, a través de los algoritmos propios del método, se realiza la estimación de un modelo matemático capaz de caracterizar las clases de los diferentes objetos teniendo en cuenta su pertenencia a dichas clases a través de la matriz de clases. Durante la etapa de validación se estudia el comportamiento del modelo estimado frente a nuevos objetos, observando la pertenencia o no a una determinada clase.



43

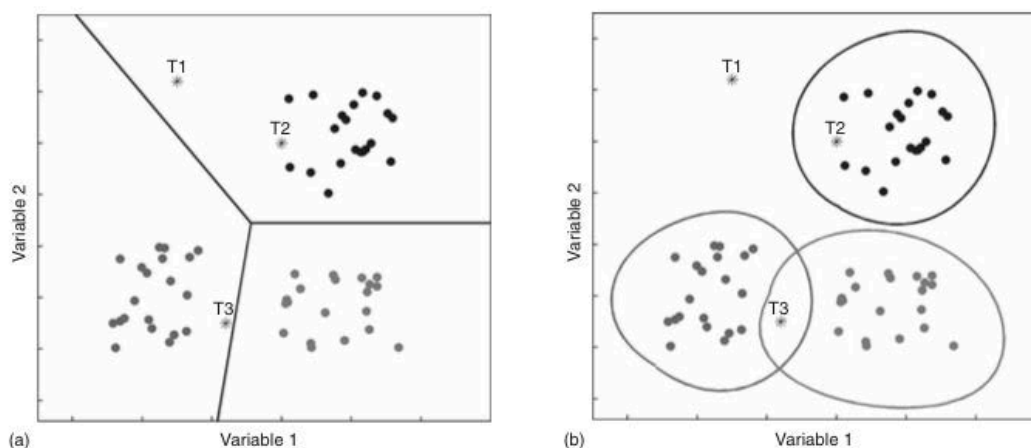


Figura 3.1: Diferencia entre modelos de clasificación (a) puros y (b) de modelado de clases

El procedimiento para la aplicación de un método quimiométrico multivariante consta de ocho etapas fundamentales. La primera de ellas es la adquisición de los datos de un conjunto representativo de muestras (objetos) a través de una o varias técnicas de análisis. La segunda etapa es la selección del método o métodos quimiométricos que se van a usar para llevar a cabo el análisis multivariante. Se trata de una etapa clave en el proceso, la elección del método utilizado marcará los resultados obtenidos por el mismo y por tanto el éxito o fracaso de la clasificación. Posteriormente, se generan las matrices de datos necesarias según el método de clasificación aplicado. Para los métodos no supervisados se genera una sola matriz \mathbf{X} (objetos) y en modelos supervisado la matriz de objetos \mathbf{X} y la matriz o vector de clases \mathbf{Y} . La siguiente etapa, de vital importancia, se trata del pretratamiento de los datos. Existen multitud de pretratamientos de datos como por ejemplo: centrado, autoescalado (varianza unidad), normalización de la señal, deconvolución, filtrado de la señal, sustracción del ruido de fondo, ponderación de las variables, aplicación de una transformación no lineal (p.e., logarítmica), derivación, etc. El efecto sobre la clasificación final de los objetos está significativamente influenciado por el pretratamiento de datos aplicado, por lo que, el tipo de



pretratamiento aplicado ha de ser estudiado y seleccionado cuidadosamente antes de la aplicación del método. Una vez hemos pre-procesado los datos y seleccionado el modelo adecuado, se realiza la estimación y optimización del modelo de clasificación. En esta etapa, en función del método de clasificación elegido, tendremos una serie de parámetros y factores que permiten optimizar el modelo a nuestro análisis en particular. Una vez obtenido el modelo, la siguiente etapa es la validación del mismo. Aquí se estudia la operatividad del mismo para asignar muestras desconocidas a una o varias clases del modelo. Se pueden realizar diferentes tipos de validación, siendo la validación interna y la validación externa las más frecuentes. Ambas se detallan en el capítulo 6.

3.3 Análisis exploratorio de datos (clasificación no supervisada)

El análisis exploratorio de datos tiene como principal objetivo la visualización sencilla de las relaciones entre muestras y variables sin necesidad de asignar las muestras a una determinada clase. Los métodos de exploración de datos son generalmente aplicados en primer lugar para tener una imagen visual global de la distribución de las muestras, observar agrupamientos, posibles objetos anómalos o discrepantes (*outliers*), relacionar muestras de similar procedencia y ayudar en la selección de los métodos apropiados para su clasificación.

El método exploratorio de datos multivariante más conocido y utilizado es el análisis de componentes principales (PCA). Sin embargo, existen otros métodos menos conocidos para llevar a cabo dicho objetivo entre los que cabe destacar el análisis de agrupamientos (*Cluster analysis*), el análisis de componentes independientes (ICA), el árbol de expansión mínima (MST) y los mapas auto organizados (SOM).

3.3.1 Análisis de componentes principales (PCA)

El análisis de componentes principales (PCA) es sin duda el método exploratorio de datos más utilizada gracias a su simplicidad y flexibilidad. El primer trabajo que define el modelo PCA data del año 1901¹⁰⁵ aunque los algoritmos usados son conocidos desde 1829 gracias a Cauchy. Entre 1930 y 1940 los métodos multivariantes cobran una mayor importancia y el PCA se empieza aplicar en todas las disciplinas como biología, geología y química¹⁰⁶. En la década de los 70 del siglo pasado el método experimenta un gran impulso y comienza a ser utilizado en química de manera rutinaria hasta nuestros días.¹⁰⁷

El método PCA lleva a cabo una transformación matemáticas sobre los datos originales que puede ser representado como:¹⁰⁸

$$\mathbf{X} = \mathbf{TP} + \mathbf{E} \quad (\text{Eq. 3.1})$$

donde:

- \mathbf{X} es la matriz de datos original con dimensiones $I \times J$, siendo las I filas el número de objetos y las J columnas las variables utilizadas en el análisis. Los elementos de la matriz representan los valores de las medidas de las diferentes muestras.
- \mathbf{T} es la matriz de puntuaciones (*scores*) que tiene las mismas filas que la matriz \mathbf{X} original y A columnas, siendo A el número de componentes principales. Por tanto, la matriz \mathbf{T} tiene dimensiones de $I \times A$ y está relacionado con los objetos de la matriz \mathbf{X} .
- \mathbf{P} es la matriz de cargas (*loadings*) de dimensiones de $A \times J$ y está relacionada con las variables de la matriz \mathbf{X} .

- Cada columna de T y cada fila de P puede ser representado como un vector columna y fila t_a y p_a , respectivamente, los cuales definen un determinado componente principal.
- El producto TP es por tanto una aproximación de los datos, donde la matriz E o de residuales representa el error.

La Fig. 3.2 representa esquemáticamente el proceso.

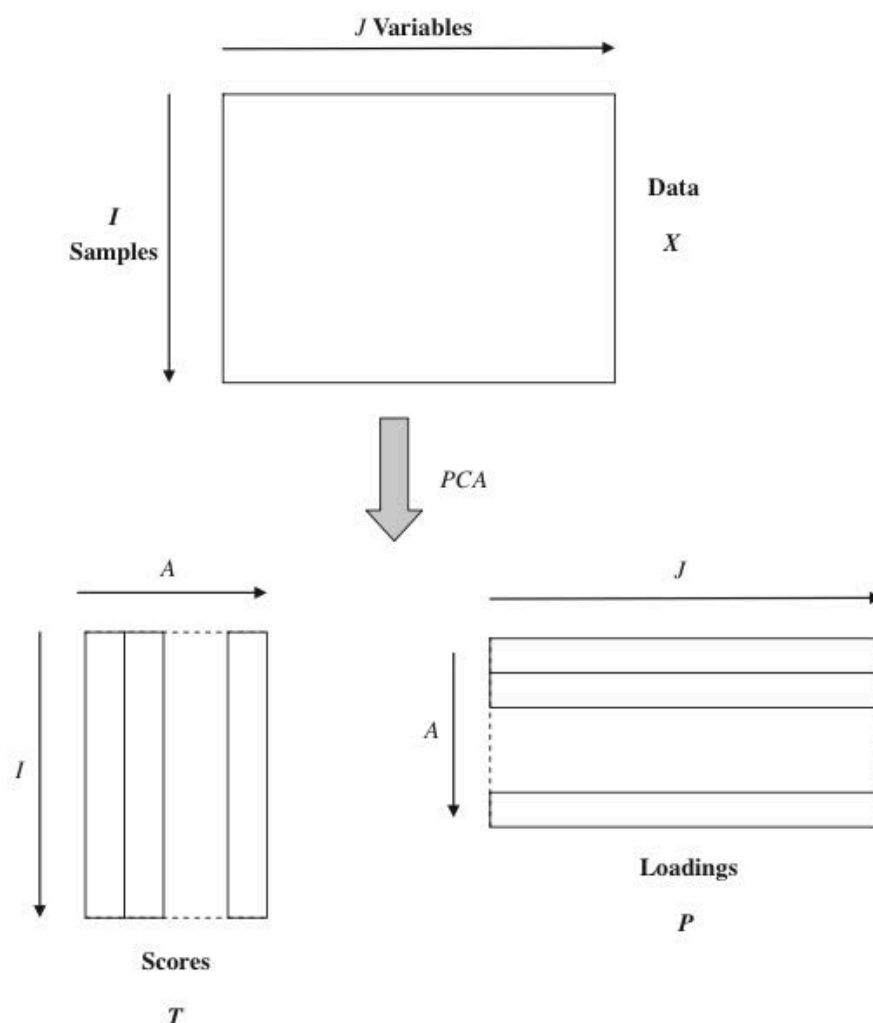


Figura 3.2: Esquema de transformación de matrices realizado por el método de PCA ¹⁰⁸

Para obtener las matrices de puntuaciones y cargas es posible aplicar diferentes algoritmos, SVD (*Singular Value Decomposition*) o NIPALS. El algoritmo SVD



obtiene simultáneamente todas las posibles componentes principales usando la matriz de correlación (\mathbf{XX}^T), donde \mathbf{X}^T es la traspuesta de la matriz de datos \mathbf{X} , mientras que el algoritmo NIPALS es un algoritmo iterativo y obtiene cada componente principal de forma secuencial.

Una de las principales ventajas de PCA es la posibilidad de realizar representaciones gráficas que proporcionan gran información sobre las propiedades de los datos a analizar. Es posible realizar la representación de uno, dos o tres vectores de puntuaciones o uno, dos y tres vectores de cargas.

Las representaciones más habituales son las *2D scores plot* y *2D loading plot*. Las representaciones de las puntuaciones nos van a dar información sobre la distribución y similitud de las muestras, mientras que las representaciones de las cargas nos da información sobre la significancia de las variables.

Los *2D scores plot* representan los scores de una componente principal (PC) frente a otra. Generalmente, el primer PC (PC1) frente al segundo PC (PC2), ya que son estos los que explican una mayor variabilidad del sistema y tienen una mayor influencia en el modelo. En ocasiones, otros PCs pueden ser necesarios para la discriminación de los objetos y por tanto es necesaria una representación tridimensional. En las gráficas *2D loading plot* se representa los *loadings* de la PC1 frente a las PC2, igual que en el caso de los scores. Estas gráficas permiten ver la relación entre variables y definir cuáles son las que tienen una mayor influencia sobre el modelo, pudiendo dar información por ejemplo sobre que partes de un espectro son útiles para discriminar entre clases.



A continuación, se describen brevemente los métodos quimiométricos lineales aplicados en la realización de esta tesis, una descripción más detallada se puede encontrar en la literatura.

El análisis discriminante es uno de los métodos de clasificación tradicionales más conocido, siendo el primer método multivariante aplicado para llevar a cabo una clasificación. Actualmente, este método está presente en la mayoría de paquetes quimiométricos comerciales y permite la realización de diferentes análisis discriminantes, entre los que cabe destacar el análisis discriminante lineal (LDA), el análisis discriminante cuadrático (QDA) y el análisis canónico de variables o análisis de correlación canónica (CVA).

En el análisis canónico de variables los objetos son separados en clases minimizando la varianza dentro de la clase y maximizando la varianza entre clases, encontrando la combinación lineal de las variables originales (direcciones) que maximizan la relación entre la varianza entre clases y la varianza dentro de la clase, y que mejor permiten discriminar las diferentes clases. Estas direcciones son llamadas funciones discriminantes y su número es igual al de clases menos uno. El QDA está basado en la regla de clasificación descrita anteriormente. La diferencia es que los límites de separación entre las diferentes clases están definidas por modelos cuadráticos y se asume que cada clase se distribuye normalmente y su dispersión (matriz de covarianza) es diferente para cada clase.

El método LDA (Fig. 3.3) es un caso especial donde las matrices de covarianza son consideradas idénticas (homocedasticidad) para todas las clases (S_P), siendo la varianza entre clases y dentro de cada clase, S_b y S_w , definidas como:¹⁰⁹

$$S_P = \frac{\sum_{i=1}^n (n_{k-1}) S_g}{k-n} \quad (\text{Eq. 3.2})$$

$$S_w = \sum_{i=1}^n \sum_{x \in c_i} (x - \mu_i)(x - \mu_i)^T \quad (\text{Eq. 3.3})$$

$$S_b = \sum_{i=1}^n n_k (\mu_i - \mu)(\mu_i - \mu)^T \quad (\text{Eq. 3.4})$$

donde n y n_k son el número de clases y el número de objetos en cada clase respectivamente, k el número total de objetos, x es cada objeto de la clase, μ_i la media de cada clase y μ es la media de todas las clases. La distancia de los objetos al centroide de cada clase se estima usando la distancia de Mahalanobis¹¹⁰, basándose en la matriz de covarianza para todo el conjunto de datos en lugar de considerar cada objeto independientemente. Al igual que CVA y QDA, LDA es un método paramétrico y por tanto asume que todas las clases tienen una distribución normal y la dispersión es la misma para todas ellas. Una limitación del LDA viene dado por la necesidad de usar como máximo un número de variables igual al número de objetos. Para ello, se lleva a cabo una reducción de variables previa, generalmente usando PCA.¹¹¹ Además, debido a que el LDA es una técnica de clasificación pura, todos los objetos son clasificados como pertenecientes a alguna de las clases descritas y, por tanto, no es posible clasificar correctamente como desconocido a muestras que no pertenecen a ninguna de las clases consideradas.⁹⁵

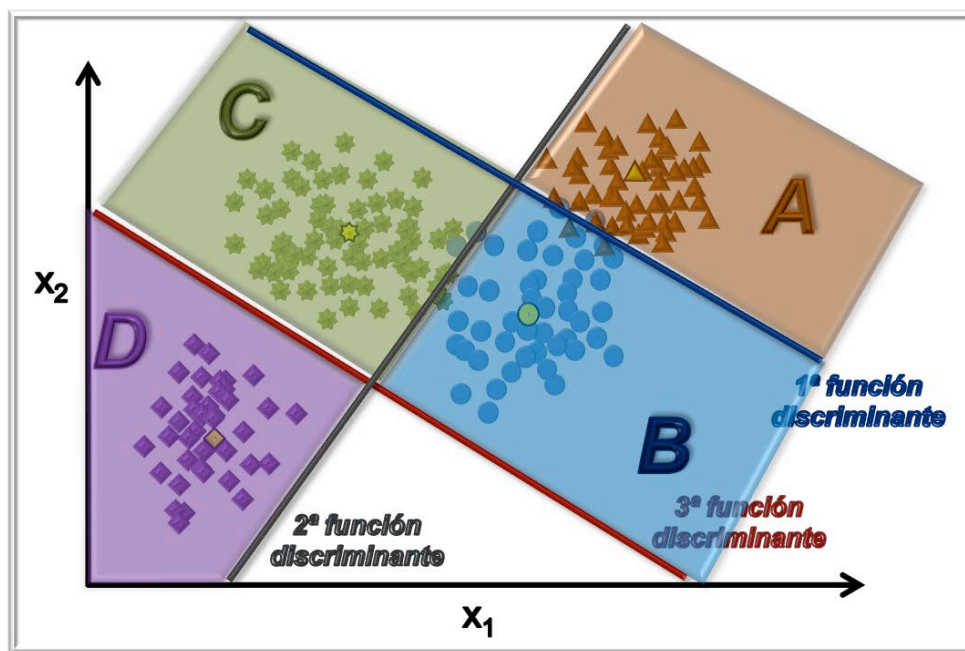


Figura 3.3: Esquema de un modelo LDA con datos simulados

3.4.2 Modelado suave independiente por analogía de clases (SIMCA)

PCA no es método útil para llevar a cabo una clasificación debido a que la información de las muestras es, en general, utilizada sólo para mostrar una visualización general de los datos. Sin embargo, los modelos de PCA pueden ser usados para llevar a cabo la clasificación a través del método SIMCA. Para ello es necesario obtener un modelo de PCA independiente para cada una de las clases. SIMCA fue uno de los primeros métodos quimiométricos aplicados en química basado en el modelado de clases. Hoy en día es, sin duda, uno de los métodos quimiométricos más importantes y extendidos en clasificación de muestras.

Como su propio nombre indica, este método es denominado suave debido a que no realiza ninguna hipótesis acerca de la distribución de variables y es denominado independiente ya que cada clase es modelada independientemente del resto.

Por lo tanto, para la aplicación del método SIMCA en la clasificación de diferentes clases, es necesario la obtención de un modelo de PCA para cada uno de ellos. La obtención del número de componentes principales significativos para cada clase se lleva a cabo a través de una validación de cruce (*cross-validation*). Con la información de cada PCA, SIMCA genera un espacio que define cada una de las clases. Cada clase es delimitada estadísticamente con un intervalo de confianza establecido a un determinado nivel de probabilidad generalmente el 95%. La varianza total residual de cada clase viene dada por:¹¹²

$$s^2 = \sum_{i=1}^n \sum_{j=1}^p \frac{e_{ij}^2}{(p-d)(d-n-1)} \quad (\text{Eq. 3.5})$$

siendo e_{ij} el valor del residual de cada elemento, p el número de variables, j el número de objetos y d el número de componentes principales considerados en el modelo PCA para cada clase.

Para la obtención de los resultados de predicción, un objeto u es proyectado sobre el espacio muestral y la distancia a cada una de las clases es comparada. Obtenida su puntuación, $t_u = x_u P$, y su valor residual, $e_u = x_u - t_u P^T$, la varianza residual de la muestra se obtiene como:¹¹²

$$s_u^2 = \sum_{j=1}^p \frac{e_{uj}^2}{(p-d)} \quad (\text{Eq. 3.6})$$

El nuevo objeto es clasificado en base a una distancia crítica establecida con la función de distribución F de Snedecor:

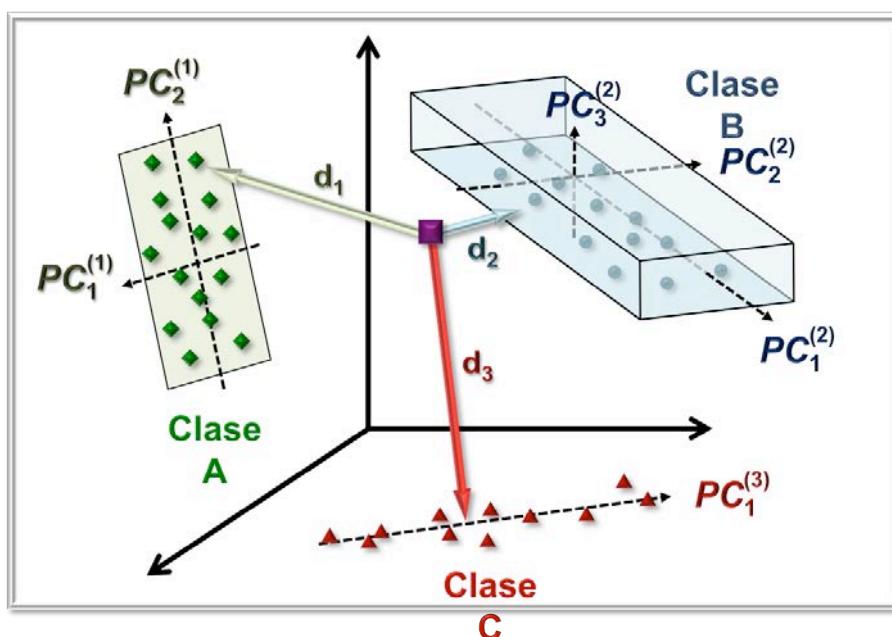


Figura 3.4: Esquema de un modelo SIMCA con datos simulados

3.4.3 Regresión Parcial de Mínimos Cuadrados – Análisis Discriminante (PLS-DA)

PLS-DA es una herramienta quimiométrica estándar presente en multitud de programas de tratamiento de datos. PLS-DA es un método de clasificación lineal basado en la regresión PLS. Dicha regresión puede ser considerada una evolución del algoritmo NIPALS, ampliamente utilizado para la obtención de modelos PCA.¹⁰⁷ El método PLS fue originalmente utilizado para llevar a cabo análisis cuantitativos (calibración multivariante). No obstante, el empleo de una matriz discriminante cualitativa como variable dependiente, permite su utilización para la estimación de modelos de clasificación. Existen dos algoritmos diferentes para llevar a cabo dicha regresión: el PLS1 cuando trabajamos con solo una variable dependiente y el PLS2 cuando hay varias variables dependientes. PLS-DA usa el algoritmo PLS2 para buscar las variables latentes que presentan una máxima covarianza entre las variables independientes (matriz X) y las variables dependientes (matriz Y).¹¹⁶ La principal ventaja del método



55

pueden obtener mediante el teorema de Bayes.¹¹⁸ El valor elegido es seleccionado minimizando el número de falsos positivos y falsos negativos y suponiendo una distribución normal de las variables.

Una parte fundamental del análisis mediante PLS-DA es la selección del número de variables latentes usadas para generar el modelo. Al igual que los modelos PCA, este número es generalmente seleccionado mediante una validación de cruce. La validación de cruce se realiza dividiendo los objetos de calibración en varios grupos de validación. Cada grupo de validación, es eliminado iterativamente del conjunto de entrenamiento o calibración. El modelo se estima con el resto de objetos y posteriormente el grupo de cancelación se introduce en el modelo para obtener su predicción. Generalmente se usan dos procesos diferentes para la división de los objetos, *venetian blinds* y bloques contiguos. El tipo de validación de cruce se selecciona en función de la estructura de los datos de entrada. Otro aspecto relevante es cuantos grupos se desea utilizar para realizar la validación. El procedimiento más común de validación cruzada es la denominada *Leave-One-Out* (LOO), en el que se selecciona el mismo número de grupos que el número de objetos, de modo que cada vez predecimos un único objeto. Sin embargo, cuando el número de objetos es elevado, el tiempo necesario para llevar a cabo dicha validación se incrementa y es conveniente reducir el número de grupos. El procedimiento LOO además, tiende a sobreestimar el poder de predicción del modelo y por tanto a elegir un número de variables latentes no adecuado. Cuando se trabaja con un número grande de clases, se recomienda el uso de un número reducido de grupos de validación con el objetivo de seleccionar un mayor número de objetos en el test y reducir la sobreestimación y mejorar la capacidad de predicción del modelo. Es frecuente usar diferentes números de grupos de validación para comparar los resultados

obtenidos, generalmente reduciendo el número de grupos y haciendo así más exigentes las condiciones de la validación.

Recientemente se ha demostrado¹¹⁶ como a través del PLS-DA es posible obtener los mismos resultados que mediante el LDA, con la mejora de la reducción de ruido y la ventaja que proporciona la posible selección de variables del método PLS. Por otra parte, al realizar el PLS-DA el modelado de las clases en función de la máxima separación de las clases, ofrece generalmente mejores resultados que los obtenidos mediante el método SIMCA.

En cuanto a la aplicación de este método al LIBS, se encuentran en la literatura varios trabajos como la clasificación de rocas terrestres y extraterrestres^{119,120}, análisis medioambiental y herencia cultural⁵⁹, análisis de suelos¹²¹ y detección de explosivos.⁴⁶

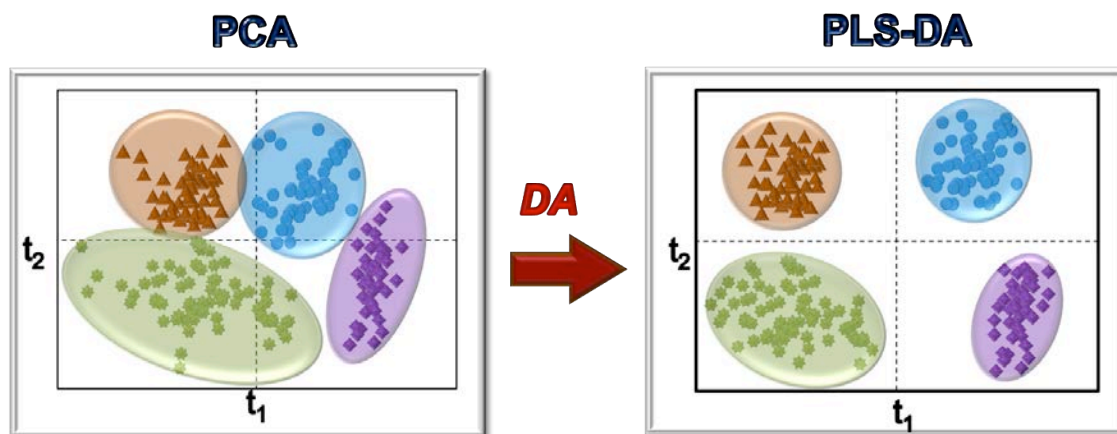


Figura 3.5: Comparación entre un modelo PCA (izquierda) y un modelo PLS-DA (derecha) con datos simulados

3.4.5 Regresión Logística Binaria (BLR)

La regresión logística está basada en la relación logística (dependencia sigmoideal) entre la probabilidad de pertenencia a una clase y las variables predictivas. En este método la decisión se toma de forma binaria. Se define una variable aleatoria binaria Z , donde $Z=1$ significa éxito y $Z=0$ error, siendo $\Pr(Z=1)=\pi$ y $\Pr(Z=0)=1-\pi$. Si tenemos n objetos, se definen las variables aleatorias Z_1, \dots, Z_n independientes con $\Pr(Z_j=1)=\pi_j$. La probabilidad total es por tanto: ¹²⁵

$$\prod_{j=1}^n \pi_j^{Z_j} (1 - \pi_j)^{1-Z_j} = \exp \left[\sum_{j=1}^n Z_j \log \left(\frac{\pi_j}{1-\pi_j} \right) + \sum_{j=1}^n \log(1 - \pi_j) \right] \quad (\text{Eq. 3.10})$$

cuando todas las π_j son iguales podemos definir:

$$Y = \sum_{j=1}^n Z_j \quad (\text{Eq. 3.11})$$

siendo Y el número de éxitos en n objetos.

Si queremos describir la proporción de éxitos en cada clase, $P_i=Y_i/n_i$, en el caso más sencillo (modelo lineal) podemos definir la probabilidad como:

$$\pi = X^T \beta \quad (\text{Eq. 3.12})$$

donde β es un vector de parámetros.

Sin embargo, el uso de una función lineal tiene limitaciones prácticas, por tanto se define el modelo general de regresión logística como: ¹²⁵

$$\text{logit } \pi_i = \log \left(\frac{\pi_i}{1-\pi_i} \right) = x_i^T \beta \quad (\text{Eq. 3.13})$$

donde x_i es un vector de medidas correspondiente a las variables ficticias.

Un desarrollo más elaborado se puede encontrar en el libro de Mc Cullagh et al. ¹²⁵



Como resumen podemos definir la probabilidad de pertenencia de un objeto a la clase 1 como P_1 y la probabilidad de pertenecer a la clase 2 como P_2 , la relación entre ambas probabilidad viene dada por la siguiente expresión:

$$\ln(p_1/p_2) = w_0 + w_1x_1 + w_2x_2 + \dots + w_jx_j \quad (\text{Eq. 3.14})$$

donde $\ln(p_1/p_2)$ se define como la transformación logística de P_1 , x_j es el valor de la variable j y w_j es el coeficiente de la variable j . La transformación logística establece una relación entre la probabilidad de pertenencia a las clases con una función lineal de las variables. Los coeficientes de la regresión logística w_j son calculados teniendo en cuenta la máxima probabilidad. Una vez los coeficientes han sido optimizados es posible llevar a cabo la predicción de muestras desconocidas.

3. 5 Métodos de clasificación supervisados no lineales

En los métodos quimiométricos supervisados no lineales, la relación entre las variables es a través de funciones no lineales más o menos complejas en función del algoritmo. Durante la realización de esta tesis dos métodos quimiométricos no lineales han sido aplicados, Máquinas de Vectores Soporte y Redes Neuronales artificiales.

3.5.1 Máquinas de Vectores Soporte (SVM)

El método SVM busca una superficie de decisión lineal o no lineal, en un espacio de variables diferentes a las iniciales que permiten separar las diferentes clases de entrenamiento con las distancias mayores entre los objetos a través de límites de decisión, es decir, los vectores soporte.¹²⁶ Los datos de los vectores de soporte establecen un límite de decisión alrededor del dominio de los datos de entrenamiento buscando la máxima separación entre clases. Cuando los datos no son linealmente

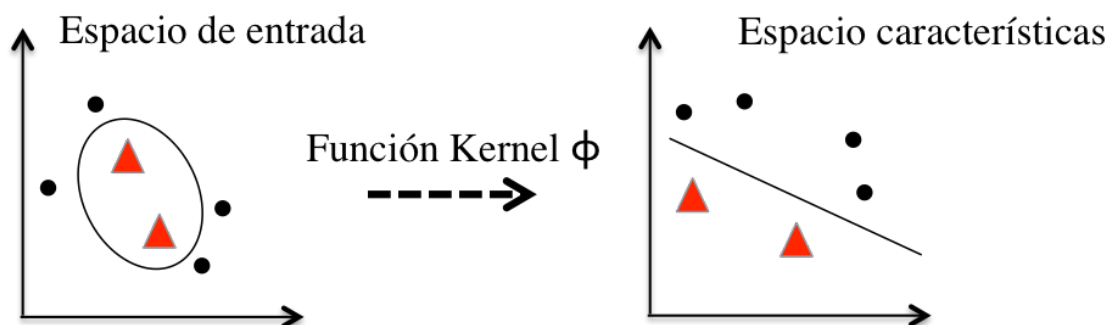


Figura 3.6: Esquema transformación Kernel en SVM

3.6 Redes neuronales

A continuación se engloba y detallan las características más importantes de los métodos basados en redes neuronales así como, sus diferentes algoritmos y arquitecturas.

3.6.1 Introducción a las redes neuronales

Las redes neuronales son una parte importante de las denominadas máquinas de aprendizaje. Las máquinas de aprendizaje son sistemas matemáticos capaces de realizar automáticamente diferentes tareas en ámbitos tales como el reconocimiento de patrones, identificación de rostros, reconocimiento por voz,¹³⁰ predicción de enfermedades,¹³¹ entre otras muchas tareas cotidianas.

Las redes neuronales han sido usadas en el conocimiento del funcionamiento del cerebro humano, cálculos simultáneos con conexiones adaptativas o para la resolución de problemas basados en algoritmos de entrenamiento, incluso cuando estos no trabajan de la misma forma que el cerebro humano.

Realizando una similitud con las neuronas humanas, podemos diferenciar tres partes principales en ellas, la zona de los axones, encargados de transportar el impulso nervioso a las sucesivas neuronas, la zona de las dendritas, encargada de recibir la

información neuronal y un cuerpo o núcleo que mantiene todas las funciones celulares de la neurona. La forma en la que las neuronas cerebrales transfieren la información entre sí es mediante la sinapsis neuronal, un aumento de la actividad en el axón produce una señal que es transferida a la neurona post-sináptica. Hay una parte del axón que genera picos eléctricos salientes provocando de la despolarización la membrana celular y con ello la transferencia del impulso. Al llegar el impulso se produce la liberación de los transmisores químicos que difunden uniéndose a los receptores moleculares de la membrana de la neurona post-sináptica.¹³² La sinapsis puede ser modulada dependiendo del número de neurotransmisores liberados y del número de receptores moleculares.

En una única capa de neuronas cada neurona recibe información del resto de neuronas. El efecto que produce la información recibida por la neurona está controlado por los pesos sinápticos, que pueden ser tanto positivos como negativos, adaptándose y permitiendo el aprendizaje para realizar una tarea o función. En el cerebro humano existen un total de 10^{11} neuronas donde cada una de las cuales se conecta con alrededor de 10^4 pesos sinápticos.¹³³ Este alto número de pesos afecta al tiempo de procesamiento haciéndolo extremadamente rápido.

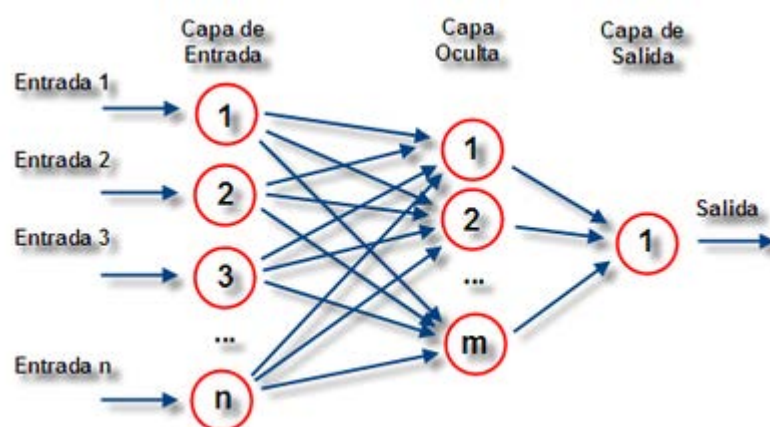


Figura 3.7: Arquitectura de una red neuronal con conexiones hacia adelante

3.6.2 Tipos de Neuronas

3.6.2.1 Neuronas lineales

Son las neuronas más simples que podemos encontrar, sin embargo presentan limitaciones computacionales. La expresión para este tipo de neuronas es:

$$y = b + \sum_i x_i w_i \quad (\text{Eq 3.16})$$

donde y es la salida de la neurona, b es un parámetro de sesgo (bias), que permite la modulación de la salida, x_i es cada entrada a la neurona y w_i el peso sináptico asociado a cada entrada o salida de la neurona.

3.6.2.2 Neuronas con límites binarios

Este tipo de neuronas (Fig. 3.8) fue introducido por Von Neumann y realizan la suma de los pesos de todos los valores de entrada hasta un valor constante cuando la suma sobrepasa un valor límite.

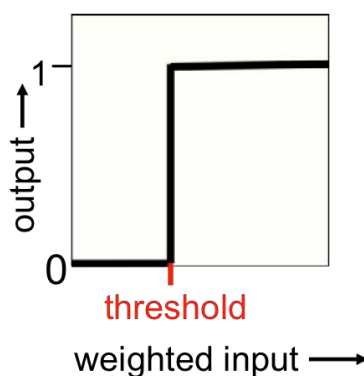


Figura 3.8: Función de transferencia binaria.

Para expresar matemáticamente dicho proceso, la suma de la neurona viene definida por la expresión $z = b + \sum_i x_i w_i$, y la salida de la neurona (y) es igual a uno si $z \geq 0$ y cero en caso contrario.

3.6.2.3 Neuronas sigmoidales

Estas neuronas producen como salida un valor real con forma de función sigmoidal acotada dependiendo de su suma. Generalmente usan funciones logísticas y tiene como principal ventaja la posibilidad de ser derivadas fácilmente, produciendo entrenamientos sencillos. Al igual que en el caso anterior la suma de la neurona se calcula mediante la expresión $z = b + \sum_i x_i w_i$ y su salida mediante la expresión $y = \frac{1}{1+e^{-z}}$, con valores continuos entre 0 y 1 (Fig. 3.9).

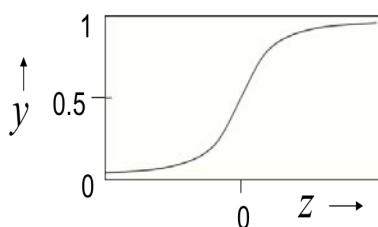


Figura 3.9: *Función de transferencia sigmoidal*

3.6.3 Entrenamiento de una red neuronal

Existen tres tipos diferentes de sistemas de entrenamiento de una red neuronal. El tipo de entrenamiento aplicado a la red neuronal tiene una gran influencia en el resultado que esta puede producir, por tanto es de suma importancia elegir el método de entrenamiento adecuado para el análisis de datos.

3.6.3.1 Entrenamiento supervisado

El entrenamiento está controlado por el usuario determinando la respuesta que debe generar la red respecto a las entradas por lo que se requiere el emparejamiento de cada vector de entrada con su correspondiente vector de salida. El entrenamiento consiste en introducir un vector de entrada, calcular la salida de la red, y determinar el error o la diferencia existente con respecto a un resultado deseado. Los pesos sinápticos se modifican de acuerdo a un algoritmo que tiende a minimizar dicho error. El



A pesar de las numerosas aplicaciones que se han hecho de este tipo de entrenamiento, son muchas también las críticas existentes sobre el mismo, ya que desde el punto de vista biológico no son intuitivos. Los algoritmos que utilizan este tipo de aprendizaje son:

- **Aprendizaje PERCEPTRON:** usado por la red PERCEPTRON
- **Regla delta o del mínimo error cuadrático:** usado por redes ADELIN y MADALIN.
- **Regla delta generalizada:** utilizada por redes multicapa.

Dentro de las redes neuronales supervisadas es importante realizar una segunda clasificación en función del objetivo del análisis. Se trata de una regresión cuando el valor de salida es un número real que está relacionado con alguna magnitud medida, como por ejemplo la concentración de un determinado analito en una muestra. O clasificación si la salida de la red es una etiqueta relacionada con una determinada clase.

3.6.3.2 Entrenamiento por refuerzo

Se trata de un aprendizaje más lento que el anterior. Se basa en la idea de no tener un ejemplo de comportamiento deseado, es decir, no tener durante el entrenamiento la salida exacta que se desea para una entrada determinada. La función del supervisor es indicar si la salida de la red se ajusta a la deseada, mediante una señal positiva (+), o no, mediante una señal negativa (-). En función de estas señales los pesos se ajustan a través de un mecanismo de probabilidades. Un ejemplo de este aprendizaje son los algoritmos denominados *Linear Reward-Penalty* y *Adaptive Heuristic Citric* que se utilizan en redes neuronales del tipo *feedforward* de tres capas.



Entrenamiento no supervisado

Durante más de 40 años este tipo de entrenamiento fue totalmente ignorado ya que era difícil definir el objetivo de este tipo de aprendizaje al no requerir la presencia de un agente externo para ajustar los pesos de las conexiones entre neuronas. Son capaces de auto-organizarse, es decir, son capaces de encontrar características, regularidades, correlaciones o categorías entre los diferentes valores de entrada.⁸ Ejemplos típicos son la regla de aprendizaje de Hebb y la regla de aprendizaje competitiva.⁹ Un ejemplo del primer método consiste en reforzar el peso que conecta dos neuronas que se activan matemáticamente de forma simultáneamente. El principal objetivo de este tipo de redes es crear una representación interna de los datos de entrada que ayude a realizar un segundo entrenamiento supervisado o por refuerzo. Es capaz de crear una representación de baja dimensionalidad de los datos de entrada, similar al proceso realizado por el método PCA. La diferencia es que en el PCA se está usando un método lineal, la red neuronal establece relaciones no lineales.

3.6.4 Tipos de asociación entre la información de entrada y de salida

Las redes neuronales son sistemas que almacenan la información del sistema modelado a través de los pesos asociados a las conexiones entre neuronas. Para ello hay que establecer una relación o asociación entre la información de los objetos introducidos a la red y la salida obtenida. Este proceso se conoce como memoria asociativa. Existen dos formas fundamentales de realizar esta asociación entrada/salida y que generan dos tipos de redes: heteroasociativas y autoasociativas.

En las redes heteroasociativas la red se entrena o calibra a partir de pares de datos (A_i, B_i) , de forma que para una determinada información de entrada A_i se asocia con su correspondiente salida B_i . En este caso la red se establece con al menos 2 capas de



3.6.5.1 Redes neuronales tipo *feedforward*

Es sin duda la red neuronal más usada en aplicaciones prácticas. Está compuesta por dos o más tipos de capas de neuronas. La primera capa (capa de entrada) es la encargada de introducir la información al núcleo de la red. Esta capa suele contener tantas neuronas como variables tengan nuestros datos. La capa salida, encargada de transferir el resultado de la red neuronal, está formada por un número de neuronas que es función del tipo de información que deseamos obtener del modelo (p.e., clasificación de objetos). Por último, entre las capas de entrada y salida se encuentran una o varias capas ocultas en función de la complejidad del sistema a modelar. Estas capas ocultas realizan una serie de transformaciones sobre los datos con el objetivo de encontrar el resultado deseado como, por ejemplo, la similitud entre diferentes objetos que permita su discriminación en diferentes clases.¹³⁶ La actividad de las neuronas en cada capa viene dada por una función no lineal de las actividades realizadas por las neuronas de la capa anterior. Es importante mencionar que en las redes *feedforward* la información fluye hacia delante y son las más sencillas en implementar. Estas redes son especialmente útiles para aplicaciones en las que no se requiera que la red retenga información de eventos pasados para evaluar eventos futuros.

3.6.5.2 Red Neuronal Recurrente

La red neuronal recurrente o completamente conectada tiene como principal característica que presenta caminos de retroalimentación entre todas sus neuronas, es decir, cada neurona está conectada a todas las neuronas de la siguiente capa, a las neuronas de la capa anterior y a ella misma a través de los denominados pesos sinápticos. Dichos pesos son modificados en la etapa de entrenamiento y son



recalculados en cada iteración (*epoch*) hasta la obtención de una red funcional capaz de realizar el proceso para el cual fue creada.¹³⁷ Se trata de redes más complejas que las redes de tipo *feedforward* ya que las neuronas de cada capa intercambian información con las neuronas de las diferentes capas. Estas redes tienen entrenamientos largos y extremadamente complejos. Sin embargo, además de ser más realistas desde el punto de vista del funcionamiento de una red neuronal biológica, son capaces de mantener una mayor cantidad de información y durante más tiempo.

Estas redes están presentes en nuestra vida cotidiana como, por ejemplo, en la predicción de un texto en un buscador de internet. Dada una secuencia de palabras estos buscadores son capaces de predecir la siguiente palabra, gracias a un entrenamiento previo intensivo.

3.6.5.3 Red neuronal simétricamente conectadas

La red neuronal simétricamente conectada es similar a la red neuronal recurrente, excepto que las conexiones entre las neuronas son simétricas, es decir, tienen los mismos pesos sinápticos entre todas las capas de neuronas en ambos sentidos. Podemos distinguir dos tipos de redes simétricamente conectadas: las que tienen una o varias capas ocultas, también llamadas máquinas de Boltzmann,¹³⁸ y las que no presentan ninguna capa oculta, también llamadas redes Hopfield.¹³⁹

El algoritmo de entrenamiento de estas redes neuronales, basado en la maximización del producto de la probabilidad asignada a los datos de entrenamiento,¹⁴⁰ es mucho más sencillo y simple que en el caso de redes neuronales recurrentes.

3.6.5.4 Red neuronal de tipo Perceptron

Esta red neuronal tiene una gran importancia histórica ya que fue la primera red neuronal con un mecanismo de entrenamiento capaz de determinar automáticamente los

pesos sinápticos para la clasificación correcta de objetos a partir de un conjunto de ejemplos o datos de entrenamiento. Fueron popularizadas por Frank Rosenblatt en 1960 gracias a su potencial de aprendizaje. En 1969 Minsky y Papert¹⁴¹ realizaron un estudio exhaustivo de este tipo de redes mostrando las limitaciones de la misma. Esta red es actualmente ampliamente utilizada en la estimación de modelos a partir de un gran número de variables.

La arquitectura de esta red está compuesta únicamente por dos capas de neuronas, una de entrada y otra de salida. La capa de entrada recibe la información y la transmite a las neuronas sin realizar ningún tipo de operación o transformación sobre las variables de entrada. En general, la información de entrada, así como la función de activación de las neuronas, es de tipo binario, como se vio en el apartado 3.6.2.2. Cada neurona de la capa de salida representa a una clase, siendo su valor igual a uno si el objeto de entrada pertenece a una determinada clase o cero en caso contrario.

Además del peso sináptico, esta red neuronal incluye otro parámetro de modulación denominado sesgo o bias (b). Este parámetro de modulación se introduce en las neuronas en la capa de entrada con un valor igual a 1 (Fig. 3.10) y tiene como función la activación de todas las neuronas de entrada.

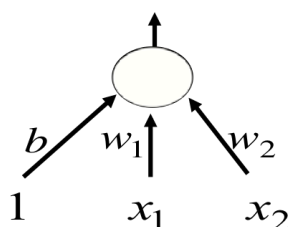


Fig 3.10: Esquema de una neurona en una red neuronal del tipo perceptron

3.6.6 Algoritmos de entrenamiento de redes neuronales

En el contexto de las redes neuronales el aprendizaje hace referencia al proceso de ajuste de los parámetros de la red.¹⁴² A partir de un conjunto de pesos sinápticos aleatorio, el proceso de aprendizaje busca aquel conjunto de pesos que permitan a la red neuronal realizar correctamente una determinada tarea. El proceso de aprendizaje es un proceso iterativo, en el cual se va actualizando la solución hasta alcanzar un nivel operacional lo suficientemente bueno.

La mayoría de los métodos de entrenamiento utilizados en las redes neuronales de tipo *feedforward* consisten en elegir una función de error capaz de establecer el estado operativo de la red en función de los pesos y estimar su error de predicción. El objetivo del entrenamiento es obtener el conjunto de pesos sinápticos que minimizan (o maximizan) dicha función. La optimización de dicha función consiste en un ajuste iterativo de los pesos en función de los datos de entrada hasta obtener los valores óptimos que hacen a la red neuronal operativa.

3.6.6.1 Método del gradiente descendiente

Es uno de los algoritmos de entrenamiento más habituales en redes sencillas. En este método se define un gradiente ($E(W)$) que indica la variación del error de la red en función de un conjunto de pesos sinápticos (W). El objetivo del aprendizaje es encontrar aquellos pesos que proporcionan un mínimo global de la función de error, aunque en muchos casos es suficiente encontrar un mínimo local. Dado un conjunto inicial de pesos, el método busca la dirección con el máximo gradiente de variación del error, y en cada iteración va modificando los pesos siguiendo dicha dirección en el sentido de máximo decrecimiento del error.¹⁴³ De este modo se va produciendo un descenso por la superficie de error hasta alcanzar un mínimo local de acuerdo a la siguiente ecuación:

$$W(t + 1) = W(t) - \alpha E(W) \quad (\text{Eq.3.17})$$

donde α representa la tasa de entrenamiento, es decir, el tamaño del paso en cada iteración. Idealmente este valor debería ser infinitesimal. No obstante, un valor demasiado pequeño puede resultar en un proceso de optimización lento y la posibilidad de acabar en un mínimo local. Por el contrario, un valor demasiado grande puede hacer que los valores oscilen en torno al mínimo global. Por tanto, la optimización de α es un punto crítico en este tipo de entrenamiento.

La velocidad del proceso de entrenamiento depende de la velocidad con la que los valores de los pesos convergen. Puede ser muy lento si dos variables de entrada están altamente correlacionadas.

3.6.6.2 Método de propagación inversa del error (*back-propagation*)

Es el tipo de entrenamiento utilizado por red tipo *feedforward* multicapa. Fue propuesto por primera vez en 1986 por Rumelhart et al.¹⁴⁴ Es un método de entrenamiento supervisado basado en el método del gradiente descendente y en el que podemos distinguir dos etapas. En una primera etapa un la información de entrada es introducido y propagado a través de las distintas capas de la red hasta generar la correspondiente salida. Esta salida se compara con la salida esperada calculándose el error cometido por cada neurona de la capa de salida. Este error es propagado hacia atrás, partiendo de la capa de salida, hacia todas las neuronas de las capas ocultas, asignando a cada una de ellas un error proporcional a su contribución al error global. Teniendo en cuenta dichos errores, los pesos sinápticos son ajustados iterativamente en cada neurona. De este modo la red evoluciona en el sentido de máxima adaptación a los datos de entrada. No es posible a priori saber cómo van a evolucionar las distintas capas ocultas, pero sí es posible medir la velocidad con la que estas cambian el error a medida

que modificamos sus pesos sinápticos. Cada neurona de las capas ocultas puede afectar a las neuronas de la capa de salida y, por lo tanto, tiene numerosos efectos independientes sobre el error que han de combinarse. Se puede calcular la derivada del error de todas las neuronas de una manera eficiente y al mismo tiempo. Una vez que obtenemos el error derivado de cada neurona, es fácil obtener el error derivado de cada uno de los pesos que van a cada neurona.

El algoritmo de este método se basa en la optimización iterativa del conjunto de pesos sinápticos que mejor relacione los datos de entrada con la salida esperada:¹⁴⁵

$$W(t + 1) = W(t) + \Delta W(t) \quad (\text{Eq. 3.18})$$

donde la variación en los pesos es proporcional al gradiente de la función de error $E(W)$:

$$W(t + 1) = W(t) + \alpha \nabla E[W(t)] \quad (\text{Eq. 3.19})$$

En la primera iteración el modelo propaga la información de entrada X_p hacia delante para obtener una salida Y_p . La salida de cada una de las neuronas que componen la red depende de su estado de activación. Considerando una función de salida identidad tenemos:

$$y_i(t) = F_i(a_i(t)) = a_i(t) \quad (\text{Eq.3.20})$$

siendo

$$a_i(t) = f_i(h_i(t)) \quad (\text{Eq.3.21})$$

La regla de propagación más simple y utilizada consiste en realizar una suma ponderada de los pesos sinápticos en cada neurona.

$$h_i(t) = \sum_j w_{ij} * x_j(t) \quad (\text{Eq.3.22})$$

Posteriormente se compara la salida obtenida Y_p con la salida deseada T_p , calculando el error de predicción e_p de cada uno de los objetos mediante la siguiente expresión:

$$e_p = \frac{1}{2} \sum_{k=1}^M (t_{pk} - y_{pk})^2 \quad (\text{Eq.3.23})$$

donde M es el número total de neuronas de la capa de salida. El error global de la red viene dado por:

$$e = \frac{\sum_{p=1}^P e_p}{P} \quad (\text{Eq.3.24})$$

siendo P el número total de objetos.

Puesto que la la modificación de los pesos sinápticos es proporcional al gradiente de la función de error (Eq.3.19):

$$\Delta w_{ji} = -\alpha \frac{\partial e_p}{\partial w_{ji}} \quad (\text{Eq.3.25})$$

Derivando la expresión 3.23 aplicando la regla de la cadena

$$\frac{\partial e_p}{\partial w_{ji}} = \frac{\partial e_p}{\partial h_j} \frac{\partial h_j}{\partial w_{ji}} \quad (\text{Eq.3.26})$$

La primera derivada describe cómo varia el error al cambiar la entrada de la neurona j , mientras que la segunda derivada describe la variación de la entrada de la neurona j al modificar el peso sináptico que conecta una neurona i con dicha neurona j . Teniendo en cuenta la ecuación 3.22

$$\frac{\partial h_j}{\partial w_{ji}} = \frac{\partial \sum_i w_{ji} y_{pi}}{\partial w_{ji}} = y_{pi} \quad (\text{Eq.3.27})$$

Denominando a la primera derivada de la expresión 3.26 como

$$\frac{\partial e_p}{\partial h_j} = -\delta_{pj} \quad (\text{Eq.3.28})$$

resulta

$$\frac{\partial e_p}{\partial w_{ji}} = -\delta_{pj} y_{pi} \quad (\text{Eq.3.29})$$

y la ecuación 3.25 se puede expresar como

$$\Delta w_{ji} = -\alpha \delta_{pj} y_{pi} \quad (\text{Eq.3.30})$$

donde δ_{pj} puede estimarse como

$$\delta_{pj} = -\frac{\partial e_p}{\partial y_{pj}} \frac{\partial y_{pj}}{\partial h_j} \quad (\text{Eq. 3.31})$$

siendo

$$\frac{\partial y_{pj}}{\partial h_j} = \frac{\partial f_j(h_j)}{\partial h_j} = f_j'(h_j) \quad (\text{Eq.3.32})$$

Para la estimación de primer término de la ecuación 3.31 es necesario diferenciar entre las neuronas de salida y las neuronas de las capas ocultas.

En el caso de las neuronas de la capa de salida, teniendo en cuenta la ecuación 3.23

$$\frac{\partial e_p}{\partial y_{pj}} = \frac{\partial \frac{1}{2} \sum_{j=1}^M (t_{pj} - y_{pj})^2}{\partial y_{pj}} = -(t_{pj} - y_{pj}) \quad (\text{Eq.3.33})$$

Combinando las expresiones 3.30, 3.32 y 3.33, la variación de los pesos de una conexión que se dirige a la capa de salida de la red puede expresarse como

$$\Delta w_{ji} = \alpha (t_{pj} - y_{pj}) f_j'(h_j) y_{pi} \quad (\text{Eq.3.34})$$

En el caso de una neurona j de la capa oculta, el primer término de la ecuación 3.31 puede expresarse como

$$\frac{\partial e_p}{\partial y_{pj}} = \sum_k \left(\frac{\partial e_p}{\partial h_k} \frac{\partial h_k}{\partial y_{pj}} \right) \quad (\text{Eq.3.35})$$

donde k hace referencia a las neuronas de la próxima capa. Teniendo en cuenta la ecuación 3.22

$$\frac{\partial e_p}{\partial y_{pj}} = \sum_k \left(\frac{\partial e_p}{\partial h_k} \frac{\partial (\sum_j w_{kj} y_{pj})}{\partial y_{pj}} \right) = \sum_k \frac{\partial e_p}{\partial h_k} w_{kj} \quad (\text{Eq.3.36})$$

y haciendo uso de la ecuación 3.28

$$\frac{\partial e_p}{\partial y_{pj}} = \sum_k -\delta_{pk} w_{kj} = -\sum_k \delta_{pj} w_{kj} \quad (\text{Eq.3.37})$$

Combinando las expresiones 3.30, 3.32 y 3.37, la variación de los pesos de una conexión que va desde una capa oculta hacia otra capa puede expresarse como

$$\Delta w_{ji} = \alpha \sum_k (\delta_{pk} w_{kj}) f_j'(h_j) y_{pi} \quad (\text{Eq.3.38})$$

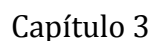
Para poder aplicar el algoritmo, es necesario elegir un valor de la tasa de aprendizaje (α). Dicho valor a de ser seleccionado cuidadosamente debido a que un valor grande puede originar oscilaciones en torno a un mínimo local o al mínimo global. Para evitar este efecto en ocasiones se introduce un parámetro en la ecuación 3.30 denominado momento (β)

$$\Delta w_{ji}(t+1) = \alpha \delta_{pj} y_{pi} + \beta w_{ji}(t) \quad (\text{Eq.3.39})$$

Con la introducción de este nuevo parámetro podemos determinar el efecto en la interacción t+1 del cambio en los pesos realizado en la iteración t, logrando llegar a una convergencia de la red con un menor número de iteraciones, ya que si la modificación de los pesos en las iteraciones t y t+1 se produce en la misma dirección y sentido, entonces el error en la iteración t+1 será inferior. En cambio, si la modificación en los pesos en las iteraciones t y t+1 se produce en direcciones opuestas, entonces el error en la iteración t+1 será superior, lo que significa que se ha pasado por un mínimo.

A modo de resumen, la siguiente tabla muestra las ecuaciones fundamentales para el método de propagación inversa del error.

$\Delta w_{ji}(t+1) = w_{ji}(t) + [\alpha \delta_{pj} y_{pi} + \beta w_{ji}(t)]$	
$\delta_{pj} \left\{ \begin{array}{l} \left(\sum_k (\delta_{pk} w_{kj}) \right) f_j' \\ (t_{pj} - y_{pj}) f_j'(h_j) \end{array} \right.$	<p style="margin: 0;">Si j es una neurona de la capa oculta</p> <p style="margin: 0;">Si j es una neurona de la capa de salida</p>



78



3.6.8 Limitaciones de los modelos de Red Neuronal

Adaptabilidad: El algoritmo de propagación inversa del error (*backpropagation*) debe hacer uso de una función de activación derivable por lo que la función de activación a de ser continua y derivable en todo el dominio de aplicación.¹⁴⁸ Esto impide la utilización de dicho algoritmo en otras redes de topología diferente como es el caso de las redes neuronales del tipo *Adaline/Madaline* donde la función de activación es discontinua.¹⁴⁸

Dependencia de los parámetros del algoritmo: Las redes que hacen uso del algoritmo de gradiente descendente han de aplicar una tasa de aprendizaje (α) que, idealmente, debería tener un valor infinitesimal. De este modo, los pesos sinápticos se van ajustando de manera más precisa haciendo que el algoritmo converja en un mínimo. Sin embargo,



Capítulo 4:

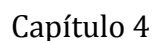
Instrumentación y montajes experimentales

4.1 Instrumentación LIBS utilizada

La instrumentación en LIBS es por lo general sencilla, de forma esquemática para la obtención de un espectro sólo es necesario el uso de un láser con suficiente potencia como fuente de excitación, diferentes lentes, espejos, rendijas o fibras ópticas para direccionar y focalizar el haz láser sobre la superficie de la muestra, un sistema de posicionamiento de la muestra, bien sea mediante plataformas con movimiento en diferentes ejes o celdas de ablación. Cuando la potencia del láser sobrepasa el umbral de ablación del material se produce una sucesión de fenómenos que llevan a la formación de un plasma como son la absorción, vaporización, atomización, excitación, formación del plasma y por último la recombinación de especie dentro del plasma. Es por tanto necesario el uso de otro sistema óptico para coleccionar la emisión del plasma inducido, compuesto de lentes, espejos, fibras. La emisión es redirigida al sistema de separación y detección de las diferentes longitudes de onda. Dichos sistemas, llamados espectrómetros, están compuestos por una rendija de entrada, un prisma o red de difracción y una cámara a modo de detector.

4.1.1 Láseres

El uso de la radiación láser en espectroscopia ha supuesto un importante avance gracias a las características únicas de esta radiación.¹⁵¹ La intensidad de dicha radiación supera en varios órdenes de magnitud la emisión por parte de lámparas, además de producir una alta monocromaticidad.¹⁵² Por tanto, antiguos experimentos difícilmente



En el mercado existe una amplia gama de posibilidades en cuanto a tipos de láseres que pueden ser usados para llevar a cabo un experimento LIBS. En la literatura nos encontramos mayoritariamente con el uso láseres de estado sólido y en particular los láseres pulsado de Nd:YAG son los más usados por la mayoría de autores. Estos láseres cuentan con un medio activo formado por un cristal de granate de Ytrio ($\text{Y}_3\text{Al}_5\text{O}_{12}$) dopado con iones Nd^{+3} , aunque cada vez se hace más frecuente el uso de láseres de $\text{Ti}:\text{Al}_2\text{O}_3$ con pulsos ultracortos de femtosegundo para diversas aplicaciones. Es posible el uso de otro tipo de láseres en LIBS, como láseres de CO_2 o incluso láseres

de diodo, sin embargo, debido a las características que proporcionan los láseres de Nd:YAG como son su estabilidad, potencias de disparo, consumo de recursos y su precio, predominan ampliamente sobre el resto. En nuestro caso, se utilizó para la realización de esta tesis un láser de Nd:YAG con Q-switch activo, modelo Brio (Quantel) trabajando en su longitud de onda fundamental a 1.064nm, con un ancho de pulso de 4ns, un diámetro del haz de 4mm y una divergencia de 0.6mrad. Además se han realizado varias colaboración con el grupo de J.J. Camacho de la Universidad Autónoma de Madrid y L. Díaz del instituto de la materia del CSIC usando un láser TEA-CO₂ para el análisis espectroscópico resulto en tiempo y espacio.^{155, 156} El láser TEA:CO₂ (Lumonics modelo K-103) trabaja con una mezcla 8:8:84 de CO₂:N₂:He. El láser cuenta con una óptica frontal multimodo de Ge (35% de reflectividad) y una red de difracción posterior de 135 líneas por milímetro optimizada para 10.6 μm. Las líneas más empleadas fueron la 9P(28), de 9.621 μm, y la 10P(20), de 10.591 μm. La divergencia del haz fue de 3 mrad.

Láseres de Nd:YAG

Se trata de un láser de 4 niveles en el que el sistema de bombeo se realiza a través de una lámpara flash de Xe que opera entre 800-820nm. La principal línea de emisión del láser de Nd:YAG se encuentra a una longitud de onda de 1064.14nm (Fig. 4.1). Además de su longitud de onda fundamental, estos láseres pueden operar a otras longitudes de onda a través de dobladores de frecuencia, pudiendo generar su segundo armónico 532nm, tercero 355nm y hasta cuarto a 266nm.¹⁵⁷ Por tanto, existe la posibilidad de trabajar en la región del IR, visible o UV según las necesidades de la aplicación. La energía de salida varía en función del sistema de generación de pulsos utilizado, si usamos el modo de régimen libre (free running) es posible generar pulsos

del orden de 100-1000 μ s con energías entre 0.1 y 100 J, consiguiendo potencias de salida de hasta 100 kW. Si la generación de pulsos es a través de un sistema Q-switch, se obtienen pulsos de entre 4 y 20 nanosegundos con energías pico de 0.01 y 1 J, lo que produce potencias del orden de 100MW.

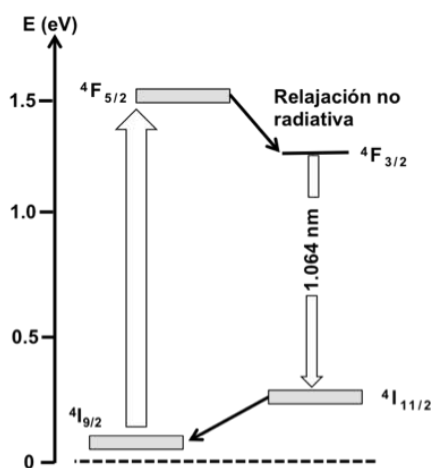


Figura 4.1: Esquema de los niveles energéticos de un láser Nd:YAG (Izq) y láser Brio (Quantel) empleado en la tesis.

4.1.2. Posicionamiento de la muestra

Sin duda, una de las principales ventajas de la técnica LIBS es la posibilidad de trabajar con muestras de pequeño y gran tamaño sin la necesidad de largas etapas de preparación de muestra o incluso el análisis directo. En nuestro caso sólo se ha trabajado con muestras sólidas y líquidas, por lo que el sistema de posicionamiento de la muestra se llevó a cabo a través de una plataforma con movimiento en los ejes XYZ con un desplazamiento de 3 μ m en cada paso. Con ello se obtiene una resolución lateral aceptable, permitiendo controlar la posición irradiada y las condiciones de focalización asegurando una fluencia similar en cada pulso, considerando las posibles irregularidades de la muestra. En todos nuestros experimentos las medidas se han llevado a presión

atmosférica y la muestra se posicionó directamente sobre la plataforma de medida sin ninguna otra consideración. En el caso de muestras líquidas, se consideraron diferentes enfoques, en primer lugar (Aceites) se midió directamente sobre el líquido, en el caso de muestras de vino, se congeló la muestra como primera opción, sin embargo finalmente se desechó y se procedió a la generación de geles sólidos mejorando las condiciones de medida. Por último en el caso de muestras lácteas, se procedió a la liofilización de las muestras. Todos los detalles son comentados en el siguiente capítulo.

4.1.3 Sistemas colectores de la radiación empleados

La interacción del haz láser focalizado sobre la muestra genera un plasma, la luz emitida por dicho plasma ha de ser recolectada y dirigida hasta el sistema de análisis de la radiación de la forma más eficaz posible evitando pérdidas de intensidad. Existen varias estrategias experimentales para llevar a cabo esta operación, las más habituales son: 1) uso de lentes que enfocan la radiación emitida por el plasma hasta la entrada del espectrómetro y 2) uso de fibras ópticas que recogen directamente la radiación de emisión. El uso de fibras ópticas simplifica el proceso de la colección de la luz, sin embargo, la luz recogida contiene información espectral limitada espacialmente y que es proporcional a la porción del volumen del plasma recogido, la cual viene dada por el ángulo descrito por la apertura numérica de la fibra. En nuestro caso el uso de fibras ópticas fue la solución elegida en la mayoría de aplicaciones. Se usaron dos fibras, la primera de ellas fue una fibra de sílice fundida con un diámetro de 1.000 μm y 1m de longitud. Por otra parte se usó otra fibra bifurcada de 200 μm y 1m de longitud. Para las aplicaciones resultas en tiempo y espacio se utilizaron lentes y espejos para coleccionar todo el volumen del plasma, permitiendo la integración espacial del mismo y con ellos su caracterización.



4.1.3 Analizadores de la radiación utilizados (espectrómetros)

El uso de espectrómetros permite la obtención de un espectro relacionando la intensidad de emisión del plasma generado por ablación con su longitud de onda. Los parámetros más importantes de cualquier espectrómetro son: el rango espectral, la resolución y tiempo mínimo de adquisición de radiación o tiempo de integración. Idealmente, nuestro espectrómetro debería ser capaz de abarcar el máximo rango espectral posible para considerar las emisiones del máximo número de elementos posibles y con una resolución suficientemente alta como para poder distinguir entre líneas de emisión muy próximas. En la actualidad existe un gran número de espectrómetros comercialmente disponibles.

Los Espectrómetros Czerny-Turner, son sin duda, los espectrómetros más habituales en LIBS. En ellos la radiación proveniente del plasma pasa por una rendija y es colimada sobre un espejo, dirigiéndose hacia una red de difracción que permite la separación de las diferentes longitudes de onda de la luz. Una vez la luz es separada en función de sus diferentes longitudes de onda, la radiación es dirigida a un segundo espejo que proyecta la radiación hacia la rendija de salida y después hacia el detector.¹⁵⁸

El uso de redes de difracción con muchas líneas permite altas resoluciones, pero disminuye el rango espectral del espectrómetro. Por el contrario redes de difracción con pocas líneas permiten rangos espectrales amplios con menor resolución. En la [Fig. 4.2](#) se muestra un esquema de un espectrómetro Czerny-Turner.

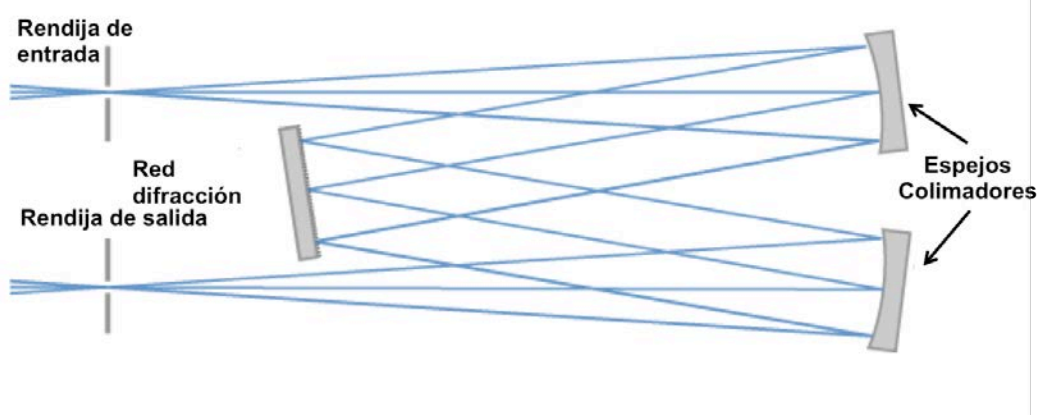


Figura 4.2: Esquema espectrometro Czerny Turner

En nuestro caso este tipo de espectrómetro se utilizó para el estudio fundamental del plasma o análisis tipo “*imaging*”¹⁵⁶ en particular se usó un espectrómetro Andor Shamrock (modelo 500i). Cuenta con una rendija ajustable y una red de difracción seleccionable entre (1200, 1800 y 2400 líneas por milímetro), que permite trabajar en la región espectral entre los 200 y 1800 nm con una resolución máxima de 0.043 nm en el primer orden. Realizando una modificación de la configuración de los espejos internos del espectrómetro podemos obtener información espacial del plasma y así observar la formación, expansión y extinción de especies y líneas dentro del plasma a tiempo real. El uso de este tipo de espectrómetros permite no solo conocer la longitud de onda e intensidad de emisión de las especies en el plasma, sino que además es posible visualizar su posición espacial dentro del plasma.

Espectrómetros Echelle

Los espectrómetros Echelle¹⁵⁹ han supuesto un gran avance en la espectroscopia debido a su gran versatilidad, permitiendo trabajar con un rango espectral entre 180-1000 nm, con resoluciones mayores a los 0.1nm. Son especialmente útiles en situaciones donde no conocemos la muestra a analizar y por tanto es necesario obtener

la máxima información espectral de la misma. Una ventaja importante de estos espectrómetros es que no presentan partes móviles lo que facilita su implementación en sistemas de campo y a nivel industrial. Sin embargo, su precio suele ser superior a los espectrómetros Czerny-Turner y es necesario el uso de lámparas para su calibración diaria. Su configuración más habitual está representada en la Fig. 4.3. Consiste en una serie de espejos y lentes de corrección para colectar y dirigir la radiación a una red de difracción con diferentes ranuras con anchuras de tamaño similar a la longitud de onda de la luz. La luz incidente es difractada y un segundo elemento dispersivo perpendicular a la red Échele actúa como un "separador de orden" o "dispersador cruzado" en la trayectoria del haz. Este tipo de espectrómetro se utilizó durante la estancia en el grupo del Prof. Pavel Veis de Bratislava (Eslovaquia) para el análisis de bronce arqueológicos.

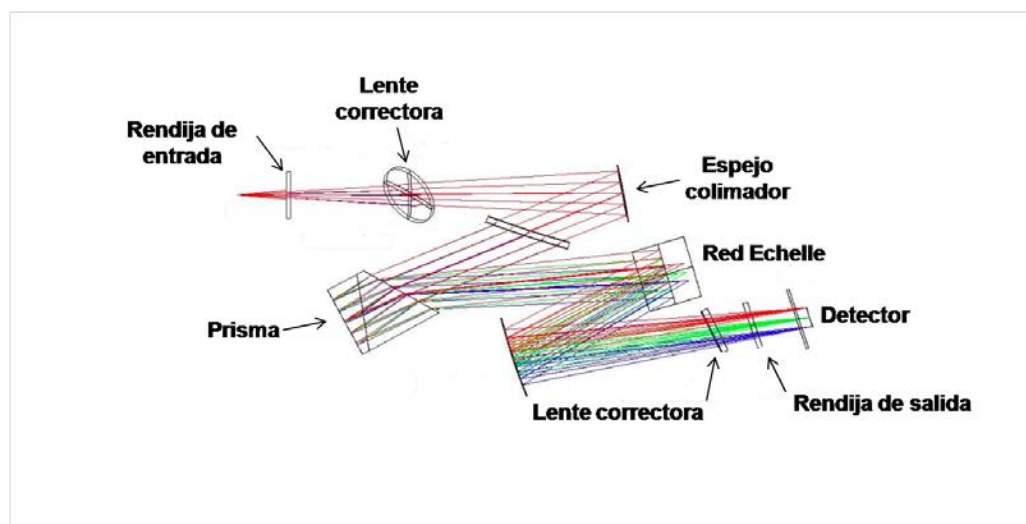


Figura 4.3: Esquema espectrómetro Échele Convencional

Espectrómetros compactos

Debido a la necesidad de generar equipos capaces de llevar a cabo medidas en campo, o análisis de muestras que no pueden ser transportadas fácilmente al laboratorio, como por ejemplo material radioactivo o explosivo, la miniaturización de los sistemas LIBS tiene una gran importancia para el desarrollo de sistema portables. Los espectrómetros han de ser por tanto robustos, portátiles y con un consumo bajo de electricidad. Entre los espectrómetros compactos destacan sobre el resto los de tipo Czerny-Turner con haz cruzado. Estos espectrómetros poseen una red de difracción fija y permiten obtener un espectro en un determinado rango espectral. El rango espectral y la resolución del espectrómetro se seleccionan en función de la necesidad modificando la red de difracción. Gracias a su tamaño y debido a que no posee partes móviles estos espectrómetros son ideales para ser incluidos en sistemas LIBS portables. En nuestro caso, la mayor parte de esta tesis se ha realizado haciendo uso de este tipo de espectrómetro. Se utilizaron dos, en primer lugar el espectrómetro utilizado fue un EPP2000, StellarNet, Tampa, FL, USA, consta de un detector lineal CCD (2048 píxeles), con una red de 300 L/mm y una rendija de entrada de 7 μm , la resolución espectral obtenida es de 0.5nm con una ventana espectral de 200 a 1.000nm. En el último año se utilizó un espectrómetro compacto (AvaSpec Dual-Channel Fiber-optic Spectrometer, Starline) de la casa comercial AVANTES (Fig. 4.4) que cubre simultáneamente el intervalo espectral entre 190 y 415 nm, con una resolución de 0.1 nm, y el intervalo comprendido entre 390 y 900 nm, con 0.3 nm de resolución espectral con una rendija de entrada de 10 μm en ambos casos y dos detectores CCD con 3648 píxeles.

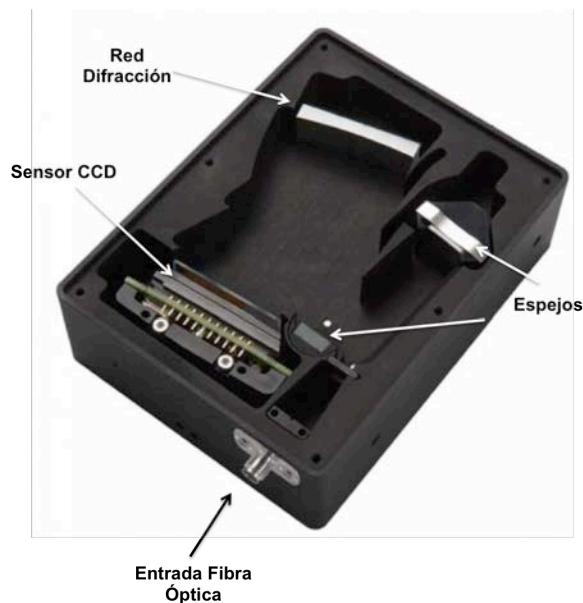


Figura 4.4: Espectrómetro compacto Czerny Turner AVANTES.

Rango espectral 200-1100 nm, resolución menor a 0.4 nm.

4.1.4 Detectores de la radiación

La función del detector o cámara es la de medir la intensidad de la radiación a cada longitud de onda previamente separada por el espectrómetro. Existen muchos sistemas para detectar y medir la intensidad de la radiación incidente, como detectores térmicos, tubos fotomultiplicadores o sistemas *array* de fotodiodos. Sin embargo todos ellos se encuentran en desuso y han sido remplazados por los sistemas electrónicos como *Charge Coupled Devices* (CCD), *Complementary Metal Oxide Semiconductors* (CMOS) y los CCD intensificados (ICCD). Muchos espectrómetros llevan el sistema de detección incluido en el interior, como es el caso de los espectrómetros compactos, sin embargo, en los espectrómetros convencionales el detector se acopla a la salida del espectrómetro permitiendo la selección del detector más apropiado para cada aplicación.

**Charge Coupled Devices (CCD)**

Estos detectores fueron desarrollados en 1969 por Willard Boyle y George E. Smith, quienes obtuvieron el premio Nobel en Física en el año 2009 “For the invention of an imaging semiconductor circuit. The CCD sensor”. Están compuestos por una estructura de píxeles lineal o en forma de matriz, con filas y columnas. Su función es acumular la carga eléctrica generada en cada pixel de cada línea o matriz proporcionalmente a la intensidad de la luz a una longitud de onda particular que incide sobre la misma. A mayor intensidad luminosa, mayor carga acumulada. Posteriormente, la carga es transformada en un voltaje, produciendo una señal analógica a su salida que es digitalizada electrónicamente produciendo un espectro o imagen. Los píxeles están compuestos por materiales semiconductores de forma que cuando un fotón índice sobre ellos, estos producen electrones que pueden ser acumulados y medidos eléctricamente. En nuestro caso los espectrómetros compactos utilizados incluían como detectores CCD lineales. De 2048 píxeles y 3648 en el caso del espectrómetro Stellarnet y Avantes, respectivamente.

Charge Coupled Devices Intensificados (ICCD)

Por otra parte y como mejora significativamente a las CCDs convencionales, se pueden multiplicar los fotones incidentes a través de un intensificador, permitiendo la obtención de espectros o imágenes en condiciones de luz extremadamente bajas o usando tiempos de exposición muy cortos del orden de ns. En esta tesis dichos detectores se usaron en las colaboraciones con J. J. Camacho y en la estancia de Bratislava. Los intensificadores están formados por tres componentes principales: un fotocátodo encargado de transformar los fotones incidentes en electrones (fotoelectrones), un disco de microcanales (MCP) que multiplica los fotoelectrones que

proviene del fotocátodo aumentando su energía cinética y por último una lámina de fósforo, que transforma de nuevo los electrones en fotones que se dirigen al sensor CCD para su detección. La eficiencia cuántica de la lámina de fósforo depende de la energía cinética de los electrones incidentes y por lo tanto de su voltaje.

4.2 Sistemas Experimentales utilizados

La mayor parte de esta tesis se ha realizado utilizando un sistema LIBS con un láser Nd:YAG de nanosegundos, sin embargo se han realizado varias colaboraciones con el grupo de J.J. Camacho y colaboradores de la Universidad Autónoma de Madrid y L. Díaz del instituto de la materia del CSIC usando un láser TEA-CO₂ para el análisis espectroscópico resuelto en tiempo y espacio.^{155, 156} Además durante la estancia en la universidad Comenius de Bratislava, bajo la supervisión de P. Veis, se utilizó un láser de Nd:YAG y un espectrómetro Echelle con un detector ICCD. Por último se realizó una estancia en el laboratorio del Prof. Vincent Motto-ros de la Universidad Claude Bernard de Lyon, utilizando un sistema LIBS compuesto de un láser de Nd:YAG y varios espectrómetros Czerny Turner con detector ICCD.

4.2.1 Sistema Experimental utilizado en la UCM

La configuración experimental utilizada con el láser de Nd:YAG se muestra en la Fig. 4.5. El sistema experimental utilizado consta de un láser de Nd:YAG con Q-switch activo, de la modelo Brio (Quantel). La focalización se llevó a cabo a través de una lente plano convexa con una distancia focal de 100mm. La emisión del plasma generado fue focalizada hasta la fibra óptica con un colimador de sílice fundida de 4mm de apertura y una distancia focal de 7 mm. La fibra óptica utilizada tiene 1.000 μm de diámetro y 1m de longitud, para dirigir la luz emitida hasta la entrada del



espectrofotómetro. La fibra óptica se posiciono a 5 cm de la superficie de la muestra y con ángulo de 60° con la normal del plano. El espectrómetro utilizado fue un EPP2000, StellarNet, Tampa, FL, USA, consta de un detector lineal CCD (2048 pixeles), con una red de 300 L/mm y una rendija de entrada de $7\ \mu\text{m}$, la resolución espectral obtenida es de 0.5 nm con una ventana espectral de 200 a 1.000 nm. Para evitar la radiación producida en los primeros instantes del plasma, bremsstrahlung inverso, se añadió al sistema un generador de retardo digital, marca Stanford modelo DG535. La muestra se colocó sobre un posicionador micrométrico XYZ manual con un paso de desplazamiento para los tres ejes de $3\ \mu\text{m}$.

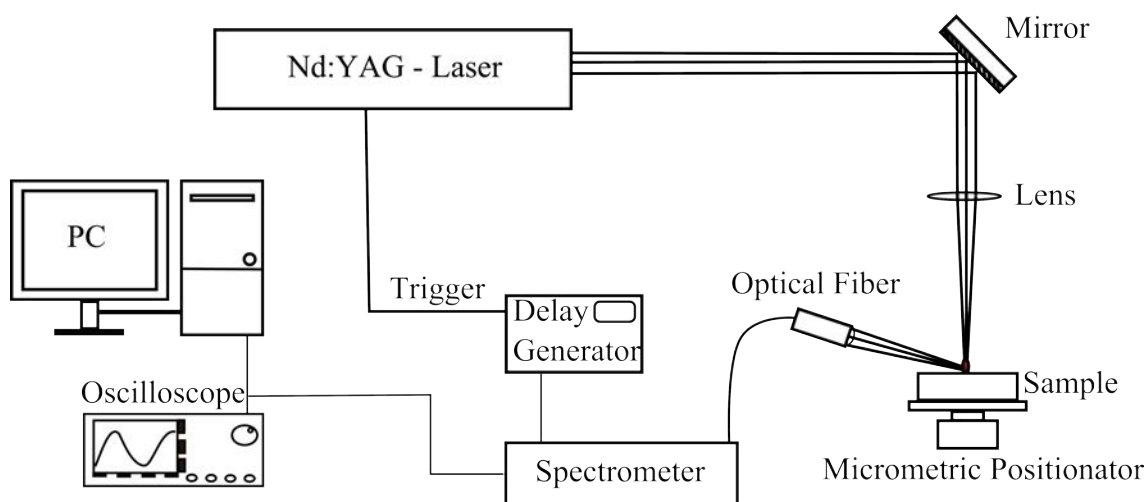


Figura 4.5: Esquema del montaje experimental utilizando un láser de Nd:YAG (arriba), foto del sistema utilizado (abajo)

4.2.2 Sistema experimental utilizado en la UAM

Se trabajó con un láser gaseoso CO₂ con una óptica frontal multimodo de Ge (35% de reflectividad) y una red de difracción posterior de 135 líneas por milímetro optimizada para 10.6 μm . Las líneas más empleadas fueron la 9P(28), de 9.621 μm , y la 10P(20), de 10.591 μm . La divergencia del haz fue de 3 mrad. El perfil del pulso se monitorizo y se observa un pulso de una anchura de 64 ns (considerando el 90 % de la energía total) y una cola de aproximadamente 3 μs , lo que produce una sobreexcitación



del plasma.

La Fig. 4.6 muestra un diagrama esquemático del sistema experimental utilizando el láser TEA:CO₂. La focalización del haz se realizó con lentes de NaCl de 24 o 40 cm de distancia focal. El spot generado tiene un radio aproximado en torno a 500 μm . Para optimizar la energía del láser utilizado se emplearon un conjunto de atenuadores de CaF₂. Para realizar los estudios resueltos en tiempo, se dividió el haz laser a través de un divisor de haz y se monitorizo a través de un fotodiodo a través de un osciloscopio (Tektronic TDS 7415). En cuanto al análisis de imágenes del plasma, se introdujo un prisma Dove de cuarzo para rotar 90° la muestra y centrarla sobre la rendija del espectrómetro, siguiendo la expansión en el eje Z, la red de difracción del espectrómetro se fijó en orden cero (espejo) para proyectar la luz del plasma directamente sobre el detector ICCS (Andor iStar DH.734, 1024x1024 pixeles, 13 μm pixeles). El retardo fue controlado directamente por el propio software de la cámara.

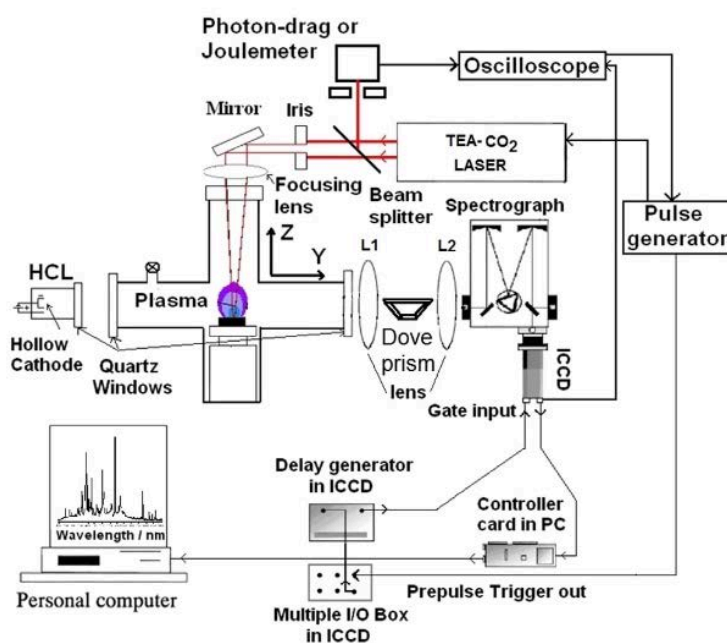


Figura 4.6: Esquema del montaje experimental utilizando un láser de TEA:CO₂ para el estudio de espectroscopia resuelta en tiempo y espacio.

4.2.2 Sistema experimental utilizado en la Universidad Comenius (Eslovaquia)

La Fig. 4.7, muestra un esquema del sistema experimental utilizado durante la estancia en Bratislava. Como resumen, se utilizó un láser Nd: YAG con un sistema de Q-switch activo y operando en su segundo armónico, 532 nm, produciendo pulsos de 4 ns, con una frecuencia máxima de repetición de 10 Hz y 165 mJ de energía máxima por pulso. El haz de láser se focalizó sobre la muestra utilizando una lente de cuarzo con una distancia focal de 40 mm produciendo un cráter de 0.1 mm de diámetro. La emisión del plasma se colectó mediante un sistema de espejos y se focalizó sobre una fibra óptica directamente acoplada a la rendija de entrada de un espectrómetro échele (Mechelle ME 5000, Andor Technology) cubriendo un rango de longitudes de onda entre los 200 y 950 nm. Como sistema detector de la radiación, se utilizó una cámara CCD intensificada, ICCD, (iStar, Andor Technology). La calibración en longitudes de onda del espectrómetro Échele, se realizó mediante una lámpara de calibración de mercurio argón (Ocean Optics, HG-1). Para más detalles se puede consultar el trabajo del Prof. Pavel Veis y colaboradores.^{160,161}

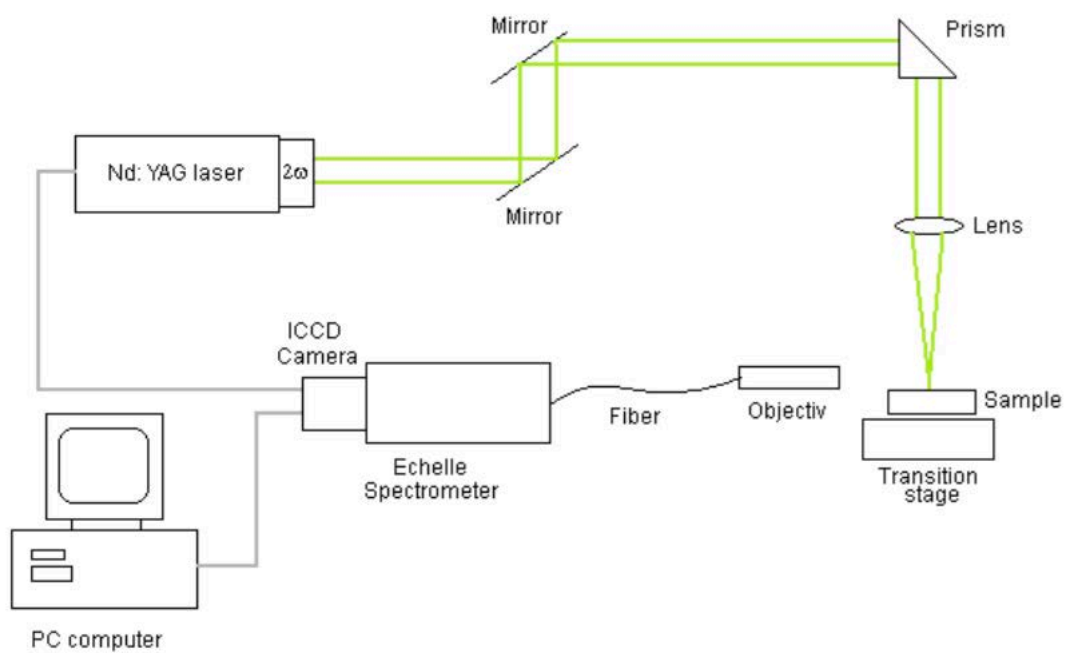


Figura 4.7: Esquema del montaje experimental utilizado en la estancia en Bratislava para el análisis de Bronces arqueológicos.(Figura propiedad de P. Vais)

Capítulo 5:

LIBS y métodos quimiométricos aplicados al análisis alimentario

En este capítulo se detalla la aplicación de la técnica LIBS en el desarrollo de metodologías de análisis para el control de calidad de muestras alimentarias. La industria agroalimentaria representa el 14.6 % de la facturación total de la unión europea con un valor superior a los 1.048.000 millones de euros. En España, la industria de alimentación y bebidas es el primer sector industrial, según la última encuesta industrial de empresas del INE (Diciembre de 2013), representando el 20.6 % de las ventas netas de productos, el 18.2 % de personas ocupadas, el 16.8 % de las inversiones en activos materiales y el 15.3 % del valor añadido.

En cuanto a la seguridad alimentaria, los consumidores esperan que el consumo de alimentos sea seguro. Los alimentos que no son seguros pueden causar efectos no deseados que en el mejor de los casos pueden ser desagradables, desde el punto de vista organoléptico y en el peor pueden causar enfermedades, generar graves secuelas o incluso producir la muerte.¹⁶² La seguridad alimentaria está íntimamente ligada a los peligros físicos, químicos y/o microbiológicos que puede haber en cualquier punto de la cadena alimentaria, que va desde la explotación agrícola o ganadera al consumo del producto. Las empresas alimentarias deben desempeñar un papel importante en el control de estos peligros. Para ello se les exige que adopten medidas preventivas, identificando y controlando los riesgos antes de poner en riesgo la seguridad de los consumidores. Con el objetivo de identificar y controlar los posibles puntos de riesgo en las cadenas de producción con el fin de asegurar la calidad alimentaria. El sistema de



Análisis de Peligros y Puntos de Control Crítico (APPCC) es el sistema de mayor reconocimiento internacional para garantizar la seguridad de los alimentos desde un enfoque preventivo. Los estándares de seguridad alimentaria vienen definidos por las normas ISO 22000 gestión de calidad ISO 9001 y la norma UNE-EN-45011. Dichos procedimientos vienen definidos por largos y costosos procedimientos analíticos bien establecidos. Además los análisis son lentos y suponen un importante gasto para las empresas del sector. Es por ello que durante la realización de esta tesis, se ha propuesto la utilización de la técnica LIBS, que gracias a sus ventajas, permite el desarrollo de procedimientos analíticos capaces de realizar el análisis de productos agroalimentarios disminuyendo los costes tanto temporales como económicos. El objetivo principal es la creación de un sistema de medida basado en la combinación de LIBS y métodos quimiométricos para la realización de un análisis, rápido, económico, fiable y robusto de muestras agroalimentarias, mejorando los sistemas de control de calidad actuales.

Los resultados comentados en esta sección han sido publicados en la revista *Applied Spectroscopy* 67 (9) 1064-72 (2013) con el título “*Application of Laser-Induced Breakdown Spectroscopy (LIBS) and Neural Networks to Olive Oils Analysis*”, en el caso de muestras de aceites y ha dado lugar a la patente P201400880 titulada “Método de análisis de bebidas alcohólicas” para las muestras de vino además de una publicación que se encuentra actualmente en preparación.

Resumen

El control en la adulteración y trazabilidad de aceites de oliva es a día de hoy un serio problema para la industria del aceite de oliva. Se ha desarrollado y aplicado la técnica LIBS en combinación con redes neuronales (NN) en la identificación, control de calidad, trazabilidad y detección de adulteración de aceites de oliva. La metodología



desarrollada ha permitido la identificación instantánea de aceites, a través del análisis de muestras representativas, usando un único pulso láser y haciendo uso de un modelo de NN desarrollado para la aplicación. En el estudio se han medido muestras de aceites españoles, italianos, griegos y argentinos. Los resultados muestran una identificación correcta de las muestras con una exactitud mayor al 95 %. Se ha demostrado que un sistema de pulso único unido al análisis de datos a través de NN, es suficiente para producir la identificación de las muestras. La metodología se desarrollará para generar una aplicación capaz de realizar el análisis en tiempo real, así como el desarrollo de un sistema portable que permita llevar análisis in situ sin la necesidad de personal altamente cualificado.

5.1.1 Introducción

El control de la adulteración, junto con la autenticación y trazabilidad son los problemas más importantes a los que se enfrentan los sistemas de control de calidad de aceites de oliva. La adulteración del aceite de oliva con aceites vegetales de menor coste es sin duda la más extendida y la más difícil de detectar, debido a la similitud entre los aceites usados. El elevado precio del aceite de oliva y sus beneficios sobre la salud^{165,166} lo hacen especialmente susceptible a sufrir fraudes con el objetivo de aumentar los beneficios de las empresas productoras. El aceite de oliva virgen extra, debido su valoración a nivel internacional, así como su coste (superior al resto de aceites vegetales), es sin duda uno de los que demanda un mayor control de la adulteración por posibles mezclas con otros aceites vegetales.

La adulteración de aceites oliva virgen extra no es siempre intencionada, además de la adicción de algún adulterante de forma intencionada, durante el proceso de elaboración se puede producir una contaminación accidental. Por ello hay descritos diferentes



procedimientos analíticos en la literatura capaces de detectar la adicción de pequeñas cantidades de otros aceite de oliva de menos calidad¹⁶⁵, aceites de avellana¹⁶⁶, aceites de semillas.¹⁶⁷

Técnicas como la cromatografía de gases,¹⁶⁸ cromatografía de líquidos (HPLC) con fase inversa,^{169,170} cromatografía de gases acoplada a masas¹⁷¹ y resonancia magnética nuclear (NMR)¹⁷² permiten detectar grados de adulteración menores al 1 % en aceites de oliva virgen extra. Todas estas técnicas han sido ampliamente usadas en la autenticación, control de calidad y detección de la adulteración de aceites de consumo, sin embargo presentan ciertas desventajas tales como, el coste del análisis y los largos procesos de derivatización.

También se han desarrollado métodos espectroscópicos para detectar adulterantes, entre los que puede destacar la espectroscopía infrarroja con transformada de Fourier (FTIR),^{177,178} espectroscopía de infrarroja medio¹⁷⁹ y técnicas basadas en Raman.^{180,181} Estos métodos han sido utilizados desde hace varios años en el control de adulteración de aceites de oliva.^{182,183}

El uso de métodos quimiométricos para el procesamiento de datos espectroscópicos ha supuesto una importante mejora a la hora de generar resultados.^{179,180,184} Entre todas las metodologías desarrolladas, en la bibliografía cabe destacar el uso de la espectroscopía Raman en combinación con métodos quimiométricos como PCA, PLS y NN, para la identificación y cuantificación de adulterantes en aceites de maíz, soja, aceite de orujo de oliva, avellana entre otros.¹⁸⁴⁻¹⁸⁶ A pesar de la similitud de los espectros obtenidos por Raman, se generan diferencias en los espectros que permiten llevar a cabo una identificación inequívoca.¹⁸⁷ Sin embargo el uso de la espectroscopia Raman no es



rutinario en los laboratorios de control debido al tiempo necesario en la preparación de muestra y a problemas de reflexión de la radiación láser.^{188,189}

La espectroscopía de ablación láser (LIBS) se ha estudiado y aplicado durante años debido a sus características únicas y su potencial aplicabilidad en diversos campos.^{28,29}

LIBS se ha convertido en una potente técnica analítica debido a su capacidad para llevar a cabo análisis rápidos sobre diferentes tipos de muestras.^{30,31} LIBS permite generar una huella digital espectral característica, representativa de su composición y por tanto única para cada tipo de muestra. Usando un procedimiento de correlación basado en NN, el sistema LIBS-NN puede ser entrenado para reconocer patrones espectrales característicos de las muestras, permitiendo evaluar las similitudes entre muestras desconocidas dando lugar a su clasificación.

5.1.2 Procedimiento Experimental

El sistema experimental utilizado se ha comentado en detalle en el capítulo 4. En este caso particular, la fibra óptica se posiciona a 5 cm de la superficie de la muestra y con ángulo de 60° con la normal del plano. La distancia de trabajo que se utilizó se optimizó para el trabajo con muestras líquidas. Debido a la onda de choque generada por el pulso láser, el trabajo con muestras líquidas produce salpicaduras que ensucian la óptica y dificultan la adquisición del espectro,¹⁹⁰ por este motivo se utilizó una lente con una distancia focal de 100 mm protegida con un cristal, el cual es más fácil de mantener limpio, permitiendo una mejora en la reproducibilidad de adquisición de los espectros. Para evitar la radiación producida en los primeros instantes del plasma, brehmstrahlung inverso, se incorporó al sistema experimental un generador de retardo digital, marca Stanford modelo DG535. En el caso particular del análisis de aceites, tiempo de retardo entre el pulso láser y la adquisición del espectro fue de 5 μ s. Este tiempo se optimizó en

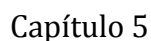


función de la relación señal ruido, tras realizar un estudio utilizando diferentes tiempos de retardo. Para mantener una energía de pulso lo más constante posible, se usó medidor de energía. La muestra se colocó sobre un posicionador micrométrico XYZ manual. El sistema puede ser dirigido a través de un ordenador personal usando una interfaz propia desarrollada en Matlab, que permite la adquisición y procesamiento de datos con la NN en tiempo real.

Se depositaron 100 mL de aceite en un vaso abierto y se realizaron las medidas LIBS focalizando directamente sobre la superficie del líquido. La viscosidad del aceite permite realizar el estudio sin ningún tratamiento de muestra previo. Después de la adquisición de las medidas no se apreciaron cambios significativos en el volumen de aceite. Las salpicaduras generalmente se producen desde la superficie del líquido en la dirección de incidencia del haz láser. Debido a la configuración experimental utilizada, no se apreciaron gotas de aceite sobre los elementos ópticos del sistema, permitiendo llevar a cabo las medidas a temperatura y presión ambiental.

Muestras de aceite

Todos los aceites utilizados fueron adquiridos en tiendas de España, Italia, Grecia y Argentina. Los aceites se seleccionaron considerando diferentes marcas y grados de calidad, incluyendo aceites de oliva, girasol, avellana y maíz producidos en diferentes países. Se usaron un total de 20 aceites, en el caso de los aceites de oliva virgen extra se seleccionaron en función de su denominación de origen incluyendo diferentes variedades como Picual (Cambril), Picual y Royal (Coosur), Hojiblanca (Hojiblanca), Picudo, Lechín, Chorrúo, Hojiblanca y Picual (Núñez del Prado) y Rotondella, Carpellese, Frantoio, Nostrale y Leccino (Pregio).



104

El algoritmo usado en el entrenamiento de la NN fue el backpropagation con gradiente descendente conjugado,¹⁹² comentado en el capítulo 3. Para detener el entrenamiento y no sobreajustar el modelo de NN se aplicó un sistema de parado previo o *early stopping*,¹⁹³ en inglés y se utilizó el error cuadrático medio (MSE) como parámetro de verificación definido en la [Eq. 5.1.1](#):

$$MSE = \frac{1}{N} \sum_k^N (r_k - y_k)^2 \quad (\text{Eq. 5.1.1})$$

donde N , y_k y r_k son el número de datos de entrada, la predicción de la NN y la salida deseada, respectivamente.

Como medida del éxito en la clasificación se ha usado el índice de coincidencia (*Spectral Correlation*).

$$SC = \frac{100}{N} \sum_i \delta_i \quad (\text{Eq. 5.1.2})$$

donde δ_i es el número de espectros correctamente clasificados y N es el número total de espectros considerados.

Cuanto mayor sea este indicador, mejor será la predicción de la NN. Por otra parte también se estudió la robustez del modelo, capacidad para clasificar como desconocidas muestras no pertenecientes a ninguna clase usada como referencia. Una mayor robustez indica una mayor capacidad para reconocer muestras desconocidas y por tanto menor tasa de muestras clasificadas incorrectamente como otro aceite.

Tabla 5.1.1 Descripción de las muestras

Muestra	Origen	Clase	Marca	Calidad	Número de replicas
1	España	1	Coosur lote 1	Aceite de oliva virgen extra	4
2	España	1	Núñez del Prado	Aceite de oliva virgen extra	3
3	España	1	Cambil	Aceite de oliva virgen extra	3
4	España	1	Hojiblanca	Aceite de oliva	3
5	España	1	Carbonel	Aceite de oliva	3
6	España	1	Ybarra	Aceite de oliva	3
7	España	1	Koipesol	Girasol aceite	3
8	España	1	La Masia	Aceite de Maíz	3
9	España	1	Coosur lote 2	Aceite de oliva virgen extra	3
10	España	1	Guinama	Avellana	3
11	Italia	2	De Marco	Aceite de oliva virgen extra	3
12	Italia	2	Pregio	Aceite de oliva virgen extra	3
13	Italia	2	Diesis-Torretta	Aceite de oliva virgen extra	3
14	Greece	3	Kretiko	Aceite de oliva virgen extra	3
15	Greece	3	KanaKis	Aceite de oliva	3
16	Argentina	4	Cocinero	Aceite de oliva extra	3



17	Argentina	4	Altavia	Aceite de oliva virgen	3
18	Argentina	4	Gretvalue	Aceite de oliva virgen	3
19	Argentina	4	Indalo	Aceite de oliva virgen	3
20	Argentina	4	Lira	Aceite de oliva virgen	3
21	Argentina	4	Nucete	Aceite de oliva virgen	3
22	Argentina	4	Oleovita	Aceite de oliva virgen	3
23	Argentina	4	Toscana	Aceite de oliva virgen	3
24	España	-	Coosur + 1% Aceite girasol	Adulterada	3
25	España	-	Coosur + 2% Aceite girasol	Adulterada	2
26	España	-	Coosur + 3% Aceite girasol	Adulterada	2
27	España	-	Coosur + 4% Aceite girasol	Adulterada	2
28	España	-	Coosur + 5% Aceite girasol	Adulterada	3
29	España	-	Cambil + 1% Aceite girasol	Adulterada	3
30	España	-	Cambil + 2% Aceite girasol	Adulterada	2
31	España	-	Cambil + 3% Aceite girasol	Adulterada	2
32	España	-	Cambil + 4% Aceite girasol	Adulterada	2
33	España	-	Cambil + 5% Aceite girasol	Adulterada	3
34	España	-	Coosur + 1% Avellana	Adulterada	3
35	España	-	Coosur + 2% Avellana	Adulterada	2
36	España	-	Coosur + 3% Avellana	Adulterada	2
37	España	-	Coosur + 4% Avellana	Adulterada	2
38	España	-	Coosur + 5% Avellana	Adulterada	3
39	Italia	-	Pregio + 1% Avellana	Adulterada	3
40	Italia	-	Pregio + 2% Avellana	Adulterada	2
41	Italia	-	Pregio + 3% Avellana	Adulterada	2
42	Italia	-	Pregio + 4% Avellana	Adulterada	2
43	Italia	-	Pregio + 5% Avellana	Adulterada	3

5.1.3 Resultados

Una vez el modelo de NN ha sido estimado, usando los espectros de referencias para todas las muestras, es posible llevar una identificación en tiempo real de sucesivos espectro de test. La Fig. 5.1.1a muestra los espectros LIBS normalizados de muestras de aceite puro: oliva (I), girasol (II), maíz (III) y avellana (IV). La composición química de cada uno de los aceites produce variaciones significativas en las líneas de emisión del espectro LIBS, permitiendo la identificación de los aceites puros. La Fig. 5.1.1b muestra los espectros normalizados de tres aceites de oliva puros de diferentes marcas y una muestra adulterada: Coosur (I), Núñez del Prado (II), Cambril (III) y Coosur + 5% girasol (IV). La composición es similar en todos los aceites de oliva, incluso cuando se ha añadido un 5% de aceite girasol, lo que hace difícil una identificación visual. En este caso debido a la rápida adquisición y teniendo en cuenta las similitudes espectrales no se realizaron promedios de espectros.

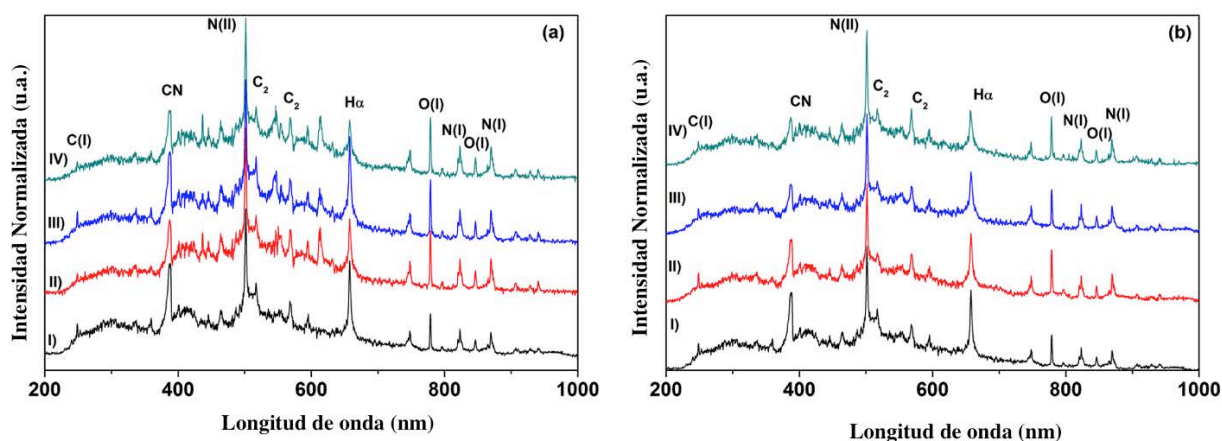


Figura 5.1.1: Espectros LIBS normalizados: (a) el aceite de oliva puro (I), (II) aceite de girasol, (III) aceite de maíz y (IV) aceite de avellana. (b) Cuatro aceites de oliva (I)Coosur lote 1, (II) Nuñez del Prado, (III)Cambril (IV)muestra adulterada (Coosur + 5% Girasol). Los espectros se han desplazado para una mejor visualización.



En trabajos anteriores se ha demostrado que los parámetros experimentales afectan al proceso global de la técnica LIBS y en especial a la intensidad del espectro, principalmente debido a las variaciones pulso a pulso.^{27, 194-195} Trabajando con un único pulso o sistema pulso a pulso, existen multitud de parámetros que afectan directamente al plasma generado y con ello al espectro obtenido. Los principales parámetros que afectan al proceso de ablación sobre líquidos son variaciones en: la interacción láser-materia, distancia focal y la energía del pulso. Además la atmosfera que rodea al plasma, presión y composición, afectan enormemente al proceso. Las zonas próximas al plasma son también excitadas y es posible observar la aparición de líneas de emisión debidas al aire.¹⁹⁶ Sin embargo, se ha demostrado que el espectro del aire no afecta al proceso de identificación de muestras, sin provocar cambios espectrales significativo.¹⁹⁷ Para tener en cuenta cómo afectan todos estos parámetros al proceso de ablación pulso a pulso, se ha representado la intensidad de emisión de la línea de H_{α} (656.279 nm) en la [Fig. 5.1.2](#). Se calculó la desviación estándar relativa obteniéndose un valor inferior al 4%, confirmando así la estabilidad y repetitividad de las condiciones experimentales utilizadas.

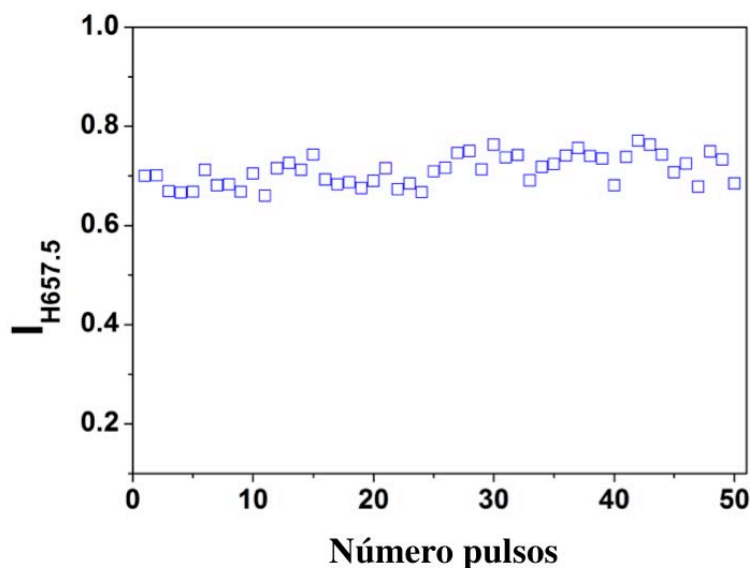


Figura 5.1.2: Variación en la emisión pulso a pulso para la línea $H\alpha$ para el aceite de oliva puro Cambril.

Primer test

El primer test realizado consiste en la evaluación del sistema LIBS/NN para identificar aceites de la misma clase que los usados en el entrenamiento. Para ello el modelo de NN se entrenó de forma binaria con los espectros de referencia de dos aceites. A continuación un nuevo conjunto de espectros (test), diferentes al entrenamiento pero pertenecientes a los aceites usados para el mismo, fueron introducidos en el modelo de NN para obtener los resultados de predicción, en lo que se conoce como validación interna. Como ejemplo, la [Fig. 5.1.3a](#) muestra gráficamente los resultados de identificación de la muestra 1 y 2. Cuando la salida de la red es 1, la muestra es identificada como perteneciente a la clase 1 y lo mismo para la clase 2. En vista a los resultados, desde el primer espectro la red es capaz de identificar el aceite correctamente y asignarlo a su clase de pertenencia. Todos los espectros excepto uno, correspondiente a la clase 1, son identificados correctamente. Teniendo en cuenta el sistema pulso a pulso utilizado, la desviación de 1 espectro de 50 es perfectamente



asumible, aunque afecta en gran medida a la correlación espectral (SC). Los resultados obtenidos para el resto de aceites se muestran en la Tabla 5.1.2, todos los aceites fueron analizados generando los correspondientes modelos de NN obteniendo una identificación correcta de todos los aceites.

La [Fig. 5.1.3b](#) muestra la gráfica ROC (*Reciver Operating Characteristic*), representando la tasa de verdaderos positivos (TPR) frente a la tasa de falso positivos (FPR), determinados como se mencionó en el capítulo 2. Además se ha obtenido el área bajo la curva (AUC) de la gráfica ROC, ampliamente utilizada para estudiar las bondades de los sistema de clasificación. Se ha observado como, sólo las diferencias espectrales entre los aceites, son las causantes de la identificación. A pesar de la presencia de líneas del aire en los espectros, se han obtenido unos valores altos en la SC para todos los aceites, indicando una correcta identificación de las muestras y por tanto permitiendo la realización de las medidas a temperatura y presión de laboratorio, sin necesidad de sistemas de vacío que dificultan y ralentizan el análisis y sin perder poder de identificación.

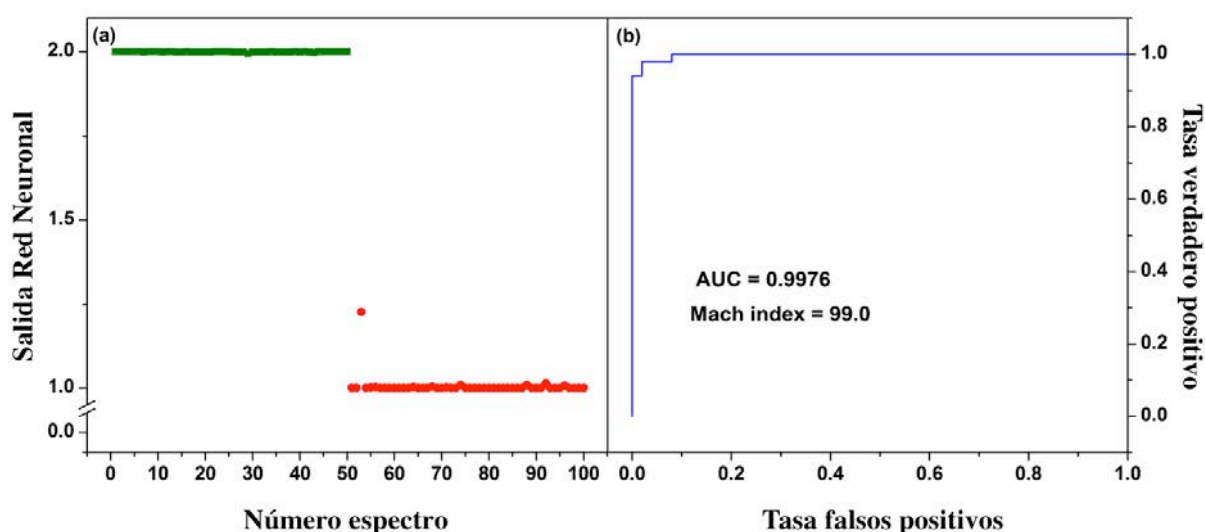


Figura 5.1.3: (a) Correlación LIBS-NN en la identificación de la clase 1. (b) Gráfica ROC para las muestras 1 y 2 (AUC = 0.9976, índice correlación= 99.0 %).

Segundo test

El segundo test realizado evalúa la capacidad del sistema en la identificación de adulteraciones en aceites de oliva. El procedimiento seguido para la realización del test es similar al caso anterior. A modo de ejemplo, la Fig. 5.1.4a muestra el análisis de aceites de oliva puros (muestra 1) y muestras adulteradas con 1, 3, y 5 % de aceite de avellana.

50 espectros de cada tipo de aceite fueron utilizados para el entrenamiento de la NN asignando cada uno de ellos con su correspondiente número de identificación, 1, 2, 3 y 4, respectivamente. Posteriormente 20 espectros no incluidos en el entrenamiento fueron introducidos en el modelo de NN para obtener su identificación. Similar al caso de aceites puros, el sistema LIBS/NN es capaz de observar diferencias entre el aceite puro y los diferentes grados de adulteración, pudiendo por tanto detectarla e incluso indicar en que grado está presente llegando a límites de tan sólo un 1 % de adulteración.

Se ha demostrado la capacidad de la NN para trabajar con más de dos muestras de referencia, lo que implica un menor tiempo de procesamiento y simplificación del método de análisis, con un tiempo aproximado de 10 s para generar el modelo de NN.

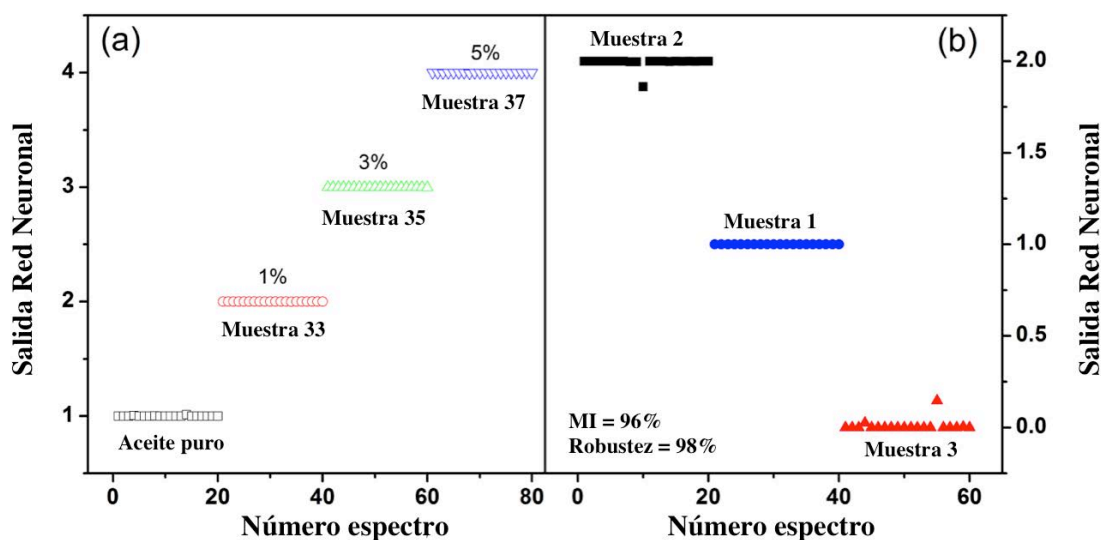


Figura 5.1.4: (a) Salida de la NN para 20 espectros de aceites de oliva COOSUR puro y adulterada con aceite de avellana, con 1, 3, y 5 %, respectivamente. (b) Salidas de la NN para la muestras 1 (Coosur), 2 (Núñez del Prado), y 3 (Cambil) desconocido para el modelo NN.

Tercer test

Por último y para completar el estudio, se ha estudiado la robustez del modelo, en otras palabras la capacidad del modelo de NN para identificar correctamente como desconocidas muestras cuyas referencias no han sido incluidas en la estimación del modelo de NN y por tanto externas e independientes al mismo. Para la realización de este último test, los espectros de la muestra 1 fueron usados como referencia en el entrenamiento de la NN y cada muestra de aceite fue eliminada del entrenamiento alternativamente, repitiendo el análisis para todos los aceites.

La red neuronal compara los espectros nuevos (test) con aquellos con los que fue entrenada, cuando coinciden, la salida de la NN es positiva y la muestra es asignada con su valor de la referencia. Si los espectros difieren significativamente con las referencias,

la salida de la NN es 0, indicando que no corresponde a ninguna clase y por tanto es desconocida para el modelo. En la [Fig. 5.1.4b](#) se muestran los resultados de un test de robustez. En este caso la muestra 1 (Coosur) y 2 (Núñez del Prado) fueron usadas para estimar el modelo de NN y la muestra 3 (Cambril) es introducida en el modelo para su predicción. Como se puede observar los espectros correspondientes a las muestras 1 y 2 son correctamente clasificados con una salida de la NN 1 y 2, respectivamente. Por otra parte, la muestra 3 es también correctamente identificada como desconocida y por tanto asignada con una salida igual a 0, indicando que no corresponde ni a la muestra 1 ni la 2 y demostrando así su robustez. Los resultados completos del último test, incluyendo las muestras de aceite puro y adulterados con un 1 %, se muestran en la Tabla 5.1.2.

Cuando nos encontramos con una muestra de aceite de oliva desconocida y sospechosa de adulteración, el analista necesita contar previamente con las muestras puras y adulteras, sin embargo en ocasiones no es posible tener acceso a dichas muestras. Este problema es común a otras técnicas analíticas que necesitan disponer de muestras de referencia. En el caso de la cromatografía, es necesario conocer el tiempo de retención de la sustancia y esto es sólo verificable con el uso de un patrón. En otras técnicas como resonancia magnética nuclear (RMN) o espectroscopia de masas el alto coste de estos patrones de análisis restringe su aplicación. Este estudio establece el origen de futuras investigaciones donde se utilicen algoritmos no supervisado complejos, tales como los algoritmos genéticos, capaces de realizar la identificación sin necesidad de usar espectros de referencia.

Es importante señalar la evidente utilidad del método desarrollado en el control de la calidad del aceite y la trazabilidad de los alimentos. Es posible realizar la identificación



in situ de los procesos de control de calidad en la industria alimentaria, obteniendo precisión adecuada y eliminando etapas de tratamiento previo de las muestras.

Tabla 5.1.2 Resultados de la clasificación de la red neuronal para aceites de oliva y muestras adulteras

<i>Correlación Espectral (%)</i>						
<i>Aceites de oliva</i>	<i>Test de identificación</i>			<i>Test de robustez</i>		
	Correctamente identificado	No identificado	Mal identificado	Correctamente identificado	Mal identificado	Desconocido
Coosur lote 1	97.4	0	0	✓	0	✓
Núñez del Prado	98.2	0	0	✓	0	✓
Cambil	98.0	0	0	✓	0	✓
Hojiblanca	99.0	0	0	✓	0	✓
Carbonel	100	0	0	✓	0	✓
Ybarra	100	0	0	✓	0	✓
Koipesol	99.2	0	0	✓	0	✓
La Masia	99.0	0	0	✓	0	✓
Coosur lote 2	96.6	0	0	✓	0	✓
De Marco	100	0	0	✓	0	✓
Pregio	99.2	0	0	✓	0	✓
Diesis-Torretta	98.4	0	0	✓	0	✓
Kretiko	96.4	0	0	✓	0	✓
KanaKis	96.8	0	0	✓	0	✓
Cocinero	100	0	0	✓	0	✓
Altavia	100	0	0	✓	0	✓
Gretvalue	98.0	0	0	✓	0	✓
Indalo	98.4	0	0	✓	0	✓
Lira	97.2	0	0	✓	0	✓
Nucete	100	0	0	✓	0	✓
Oleovita	100	0	0	✓	0	✓
Toscana	98.4	0	0	✓	0	✓
Coosur + 1% aceite girasol	97.8	0	0	✓	0	✓
Cambil + 1% aceite girasol	100	0	0	✓	0	✓
Coosur + 1% avellana	100	0	0	✓	0	✓
Pregio + 1% avellana	98.4	0	0	✓	0	✓

Número de espectros usado en el entrenamiento

El número de espectros usados en el entrenamiento de la red neuronal es un parámetro clave en la estimación de un modelo de clasificación. Es importante optimizar el número de espectros usados para el entrenamiento con el objetivo de generar un modelo con una alta capacidad de identificación de muestras pertenecientes a las clases entrenadas, como de muestras que no pertenecen a ninguna clase considerada. En la [Fig. 5.1.5](#) se muestra como varía la robustez del modelo y el valor de correlación espectral (SC) en función del número de espectros usados en el entrenamiento. Se puede ver como a medida que aumentamos el número de espectros utilizados, la robustez y el valor de la SC mejora. También se observa como incluso usando un número bajo de espectros para realizar el entrenamiento, por ejemplo 8, la robustez es aceptable. Se observa un comportamiento similar en el valor de la correlación espectral. Por tanto es importante usar, siempre que se pueda, un elevado número de espectros, al menos 50, para obtener una buena robustez y valores aceptables de correlación espectral. En nuestro caso esto no supone un problema ya que la obtención de espectros es prácticamente inmediata. Sin embargo, esto no es siempre posible en casos donde la cantidad de muestra que se dispone es pequeña, siendo necesario el uso de un menor número de espectros en detrimento de las capacidades del modelo de clasificación.

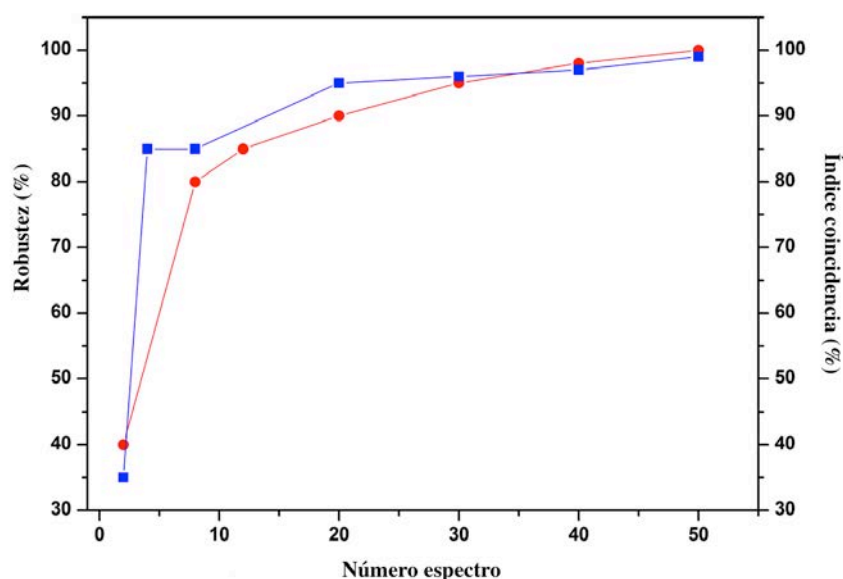


Figure 5.1.5: *Índice de correlación para la muestra 1 (rojo) y la robustez obtenido para la muestra 11 (azul) en función del número de espectros utilizados en el proceso entrenamiento.*

Se realizó un test para estimar la validez del modelo de NN en función del tiempo, para ello se realizaron nuevas medias cada 15 días durante 2 meses y se introdujeron en el primer modelo generado con las librerías de referencia del primer día, demostrando la estabilidad del sistema y del modelo a lo largo del tiempo.

No se han estudiado adulteraciones menores al 1 % debido al bajo impacto económico que supone, aunque pueden ser interesantes desde el punto de vista del potencial de la metodología. Los resultados han demostrado como un sistema de medida pulso a pulso, a pesar de presentar una menor repetitividad, es suficiente para llevar a cabo la clasificación. La intensidad de los espectros varía pulso a pulso y día a día, sin embargo la red neuronal es capaz de disminuir el efector de dichos cambios y adaptarse al reconocimiento e identificación de muestras debido a su complejidad no lineal, capaz de tener en cuenta todas estas fluctuaciones. La red neuronal es capaz de modelar, a través de la modificación de los pesos y sesgos, en el proceso de entrenamiento, las similitudes



entre espectros de la misma clase para poder realizar una generalización, pudiendo predecir el comportamiento de nuevos espectros desconocidos para el modelo con una alta tolerancia al ruido y a la presencia de espectros anómalos.

Principales conclusiones de la identificación y adulteración de aceites

Se ha demostrado como un análisis simple y preciso basado en LIBS y NN es capaz de identificar con un 100 % de confianza muestras de aceite de consumo tanto para muestras conocidas como desconocidas para el modelo. Se ha demostrado como las variaciones entre los espectros de diferentes tipos de aceite son responsables de la identificación, a pesar de la posible contribución de líneas del aire, permitiendo el análisis y reduciendo los costes sin afectar significativamente a la capacidad de discriminación del modelo.

La aplicación de LIBS/NN a aceites de oliva virgen extra procedentes de diferentes regiones de España y del mundo, incluso de diferentes años de cosecha ha producido resultados satisfactorios con un 100 % de los aceites correctamente identificados. Se han realizado diferentes test para evaluar el modelo de NN a fondo incluyendo incluso, estudios a largo plazo. Pequeños cambios en las condiciones experimentales como la variación de las intensidades del espectro debido al uso de pulsos individuales no fueron relevantes en la identificación de las muestras. Es importante destacar que para esta aplicación el uso del espectro completo como huella espectral para el entrenamiento de la red neuronal, sin realizar una reducción de variables, permite una diferenciación correcta de todos los aceites de oliva. Los test realizados al modelo verifican que la metodología es adecuada y puede producir el desarrollo de una aplicación práctica. Teniendo en cuenta la velocidad de la medida, el método desarrollado en este se podría utilizar para llevar a cabo análisis automáticos, fiables y robustos en tiempo real.

5.2 LIBS aplicado al control de fraude en Vinos

Resumen

En esta aplicación se ha desarrollado un método para el control de la adulteración y la determinación del origen geográfico de vinos con Denominación de Origen Protegida (DOP) a través de LIBS y NN. Debido a las dificultades que presenta el análisis de líquidos, se ha utilizado la gelificación y secado del gel como preparación de la muestra. El uso de un gel de colágeno permite obtener los resultados de la clasificación de una manera rápida y sin las limitaciones y dificultades de trabajar con muestras líquidas. El modelo de NN generado se ha evaluado a través de tres procedimientos de validación teniendo en cuenta su sensibilidad (validación interna), capacidad de generalización y su robustez (validación externa independiente). Los resultados muestran que la combinación LIBS/NN permite clasificar vinos en función de su DOP y detectar adulteraciones por mezcla de vinos con una metodología de alta sensibilidad, capacidad de generalización y robustez.

5.2.1 Introducción

España es uno de los grandes productores de vino del mundo siendo el primer país respecto a la superficie sembrada, el tercero en producción y el segundo mayor exportador en términos de volumen. La industria del vino es un sector muy importante en términos económicos, ambientales y sociales, así como en la imagen que proyecta España al resto de países.

El vino está constituido de dos moléculas principales (agua y etanol) y otros constituyentes mayoritarios como moléculas orgánicas y azúcares.^{198,199} Otros compuestos minoritarios son también parte del vino, lo que hace del vino una matriz compleja y difícil de analizar. Las propiedades organolépticas del vino son originadas



por pequeñas cantidades de diferentes compuestos aromáticos y taninos, que son importantes sustancias sápidas presentes en el vino tinto. La industria del vino ha sufrido un gran crecimiento en la última década debido al aumento continuo del consumo, llegando a cifras de 28 L per cápita. Sin duda, es importante controlar y asegurar las características del vino, tales como la variedad, el origen de la producción y la tecnología de elaboración del vino con el fin de preservar su calidad y procedencia. La denominación de Origen Protegida (DOP) e Indicaciones Geográficas Protegidas, (IGP) constituyen el sistema utilizado en España para diferenciar la calidad del vino en base a características propias y diferenciales, debidas al medio geográfico en el que se producen las materias primas y la elaboración del producto. La calidad de los productos vitivinícolas se determina en base al tipo de uva, su origen, año de cosecha y tipo de vendimia. Sin embargo, estos indicadores no son fácilmente reconocibles por el consumidor y la denominación de origen protegida (DOP) es una señal de calidad única e inequívoca, decisiva para los consumidores, afectando en su decisión de compra. Es necesario contar con un sistema que asegure la calidad de los vinos capaz de evitar fraudes y preservar la DOP así como las características especiales de los vinos.

Desde el año 2008, la industria vitivinícola española no solo es responsable del control del fraude en vinos, sino también de la falsificación de las botellas. Dicha legislación ha supuesto un crecimiento significativo de la industria del vino en nuestro país en los últimos años, aumentando el número de consumidores en busca de productos de calidad certificada que garanticen la procedencia y las características especiales de los productos. El fraude está presente de diversas formas en la industria vitivinícolas y produce un importante deterioro de la imagen del sector, con pérdidas valoradas en decenas de millones al año. Para evitarlo, existe una amplia legislación y un sistema de



garantía de calidad destinado a proteger la industria vitivinícola preservando las características especiales de los vinos.

Entre los fraudes más frecuentes se encuentran:

- La adición de azúcar para aumentar la graduación alcohólica, permitida en casos excepcionales y estrictamente regulado.
- La adición de agua al mosto para suplir la merma que se produce durante la fermentación.
- Falsificación del etiquetado.
- Adicción de vinos de menor calidad.
- Adición de productos químicos que mejoran las propiedades organolépticas. En algunos casos su adición está permitida hasta un nivel máximo establecido por ley.

Para el control de la DOP y la adulteración actualmente se llevan a cabo análisis sensoriales y determinaciones analíticas estándar. Este tipo de análisis no es suficiente para asegurar DOP y se requiere de técnicas cromatográficas para llevar a cabo un análisis de cada componente del vino por separado, siendo un proceso lento y costoso. Como alternativa a los métodos cromatográficos, la resonancia magnética nuclear (NMR) ha sido también usada para realizar la identificación de la variedad de las uvas,²⁰⁰ necesitando una mayor cantidad de muestra y el uso de una biblioteca espectral amplia, lo cual implica el aumento de los costes del análisis.

Existe la necesidad de desarrollar una metodología de análisis capaz de asignar las muestras de vino en función de su denominación de origen, capaz de detectar de forma sencilla, económica y rápida la adulteración y con ello evitar el fraude. La presencia de elementos minerales en el vino está generalmente relacionado con la composición del

suelo, la variedad de uva, condiciones climáticas, las levaduras de fermentación.²⁰¹ Debido a la capacidad de LIBS para producir una huella espectral de las muestras analizadas y basándose en las diferentes composiciones de los vinos analizados, se ha desarrollado una aplicación basada en LIBS y NN capaz de llevar a cabo este tipo de análisis. LIBS proporciona la huella digital de cada tipo de vino. Usando una método quimiométrico como las NN, es posible procesar de forma multivariante y de manera eficiente la información espectral, permitiendo llevar a cabo la clasificación.

Se han realizado los estudios experimentales para lograr los siguientes objetivos: (1) desarrollar un método simple, directo y rentable basado en el análisis LIBS y NN para clasificar de vinos pertenecientes a las once DOP más importantes de España, (2) evaluar la capacidad de generalización del modelo de NN en la clasificación de pertenecientes a una de los once DOP, pero no incluidos en el entrenamiento y (3) evaluar la robustez del modelo al introducir vinos totalmente independiente al modelo, vinos de mesa (sin DOP) y vinos extranjeros.

5.2.2 Sistema experimental

El sistema experimental utilizado es el mismo que el descrito en el apartado anterior. En este también se realizó un estudio para optimizar el tiempo de retardo entre el pulso láser y la adquisición del espectro fijándose en 2 μ s. Igual que en el caso de aceites, este se optimizó en función de la relación señal ruido.

Muestras de vino

Para la realización del estudio se usaron 31 vinos tintos españoles, de los cuales 24 pertenecen a una de las 11 denominaciones de origen protegida seleccionadas, tres vinos tintos extranjeros y cuatro vinos de mesa, todos ellos fueron comprados en un supermercado local. La [Fig. 5.2.1](#) muestra un mapa de España con las denominaciones

de origen utilizadas marcadas. Los vinos fueron seleccionados considerando las principales regiones españolas de vino (La Mancha, Ribera de Duero, Rioja, Valdepeñas, Vinos de Madrid, Cariñena, Ribeiro, Ribera del Guadiana, Navarra, Somontano y Toro), por otra parte un vino alemán, francés e italiano fueron también incluidos en el estudio. La mayoría de los vinos fueron elaborados con uvas de la variedad Tempranillo, Cabernet Sauvignon y Garnacha, aunque también se consideraron variedades como Tinta de Toro y Shyrah. En cuanto a la vendimia, se utilizaron sólo vinos jóvenes del año 2011. El motivo de esta elección, además de que se trató de un año con una cosecha excelente, es que, al tratarse de un vino joven, no ha sufrido de un proceso de envejeciendo en barrica y se eliminan posibles productos derivados este proceso. La Tabla 5.2.1 muestra una relación detallada de las muestras utilizadas, incluyendo su identificación, marca comercial, tipo de uva y la cosecha.



Figura 5.2.1: Mapa de las Denominaciones de Origen Protegidas de vinos españoles utilizadas.

**Tabla 5.2.1:** Vinos utilizados en el estudio

PDO	Identificador	Marca comercial	Variedad de uva
La Mancha	M1	Libertario	Tempranillo, Garnacha
	M2	Vereda Mayor	Tempranillo
	M3	Don Lucio	Tempranillo
	M4	Fidencio	Tempranillo, Garnacha
	M5	Monte don Lucio	Cabernet Sauvignon
Ribera del Duero	D1	Dehesa	Cabernet Sauvignon
	D2	Valpincia	Tempranillo
	D3	Barón de Santuy	Tempranillo, Cabernet
	D4	Mayor de Castilla	Tempranillo
	D5	Sangre de Castilla	Tempranillo
Rioja	R1	Viña Espolón	Tempranillo, Garnacha
	R2	Antaño	Tempranillo, Garnacha
	R3	Solar Viejo	Tempranillo, Garnacha
	R4	Barón de Urzande	Tempranillo, Garnacha
	R5	Castillo de Albali	Tempranillo
Valdepeñas	V1	Vega del Cega	Tempranillo
	V2	Calle Real	Tempranillo
	V3	Viña Albali	Tempranillo, Cabernet
	V4	Señorío de los Llanos	Tempranillo
	V5	Tanis	Tempranillo
Vinos de Madrid	VM1	Puerta de Alcalá	Tempranillo, Syrah
	VM2	Vega Madroño	Tempranillo, Garnacha
	VM3	Alma de Valdeguerra	Tempranillo
	VM4	Puerta de Hierro	Tempranillo
	VM5	Jesús Díaz	Tempranillo, Syrah
Cariñena	CR	Castillo de Aguaron	Garnacha, Syrah
Ribeiro	RB	Pazo	Tempranillo, Cabernet
Ribera del Guadiana	RG	5 Viñas	Tempranillo, Garnacha
Navarra	NV	Diácono	Tempranillo, Garnacha
Somontano	SM	MonteSierra	Tempranillo, Cabernet
Toro	TR	Cermeño	Tinta de Toro
Vino de mesa	TW1	Conde Noble	Mixture
	TW2	Don Simon	Mixture
	TW3	Eroski	Mixture
	TW4	Viñas Altas	Mixture
Chianti	CH	Corte Alle Mura	Sangiovese
Dornfelder	DR	Dornfield	Dornfelder
Valle Central	VC	Cimaraosa	Cabernet Sauvignon

Preparación de muestra

El análisis directo de muestras líquidas por LIBS, como se utilizó en el caso de las muestras de aceites, es difícil de llevar a cabo debido a las salpicaduras que llegan a la óptica. Debido a la menor densidad y viscosidad de las muestras de vino, las salpicaduras impedían el análisis efectivo de las muestras y esta idea fue desechada. Posteriormente se realizaron pruebas de congelación de las muestras líquidas con nitrógeno líquido para análisis por LIBS.¹⁸⁷ En el caso particular del vino, este procedimiento tampoco fue efectivo. Debido a la temperatura generada en la zona de interacción con el pulso láser, la muestra es licuada rápidamente impidiendo el análisis. Finalmente, para evitar este problema, se desarrolló un tratamiento de muestra basado en la formación de un gel de vino y su posterior solidificación y secado. Con ello no solo conseguimos tener una muestra sólida sino que también la eliminación del agua permite una preconcentración de los elementos y con ello un aumento en la intensidad de las líneas espectrales, disminuyendo los límites de detección. Para la preparación de dicho gel, se usó un colágeno comercial. Se tomaron 50 ml de la muestra de vino en un vaso de precipitados y se añadió 1 g de colágeno en polvo, el gel se disolvió y se dejó reposar 15 min hasta la formación de un gel. A continuación, las muestras se introdujeron en un horno de ventilación forzada a $35 \pm 2^{\circ}\text{C}$ durante 12 h para evaporar el exceso de agua, obteniéndose un sólido seco. El tiempo de secado fue estudiado y seleccionado en función de la relación señal ruido frente a las horas de secado ([Fig. 5.2.2](#)). Todas las muestras de vino se abrieron el mismo día para evitar la degradación y la oxidación de los constituyentes.

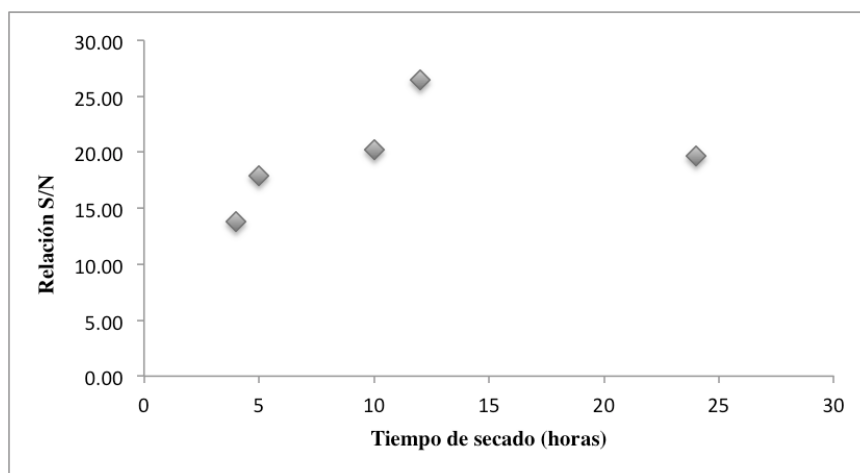


Figura 5.2.2: Estudio tiempo de secado en función de la relación Señal/Ruido considerando la línea de Ca (II) a 317.94nm y el fondo espectral

Medidas LIBS

Todas las muestras se midieron a temperatura ambiente y presión atmosférica. Cada espectro se obtuvo como promedio de tres pulsos desplazando el posicionador 0.25 mm entre pulso y pulso para exponer una nueva porción de la superficie de la muestra y evitar zonas irradiadas previamente. Para disminuir las variaciones espectrales, producidas por cambios en la energía del pulso láser, cada espectro se normalizó con la intensidad de una línea espectral de un compuesto mayoritario como K (I) 766.49 nm. Se tomaron un total de 100 de espectros para cada muestra, generando una librería espectral. Solo para las muestras M1, D1, R1, V1 y VM1 se prepararon 3 geles independientes para ser utilizados como réplicas. La réplica 1 de estas muestras se dividió en dos subgrupos, el primer subconjunto de 80 espectros se utilizó para el entrenamiento y obtención del modelo y un segundo subconjunto de 20 espectros para la validación y test. Las librerías espectrales son generadas en tan solo 2 minutos considerando la adquisición tiempo de integración del espectrómetro y la frecuencia de los pulsos de láser fijada a 1 Hz.

La reducción del número de variables ha demostrado ventajas importantes en los procesos de clasificación como la mejora de la robustez y un aumento en poder de discriminación. En este trabajo, se ha realizado una reducción del número de variables mediante la selección de intervalos espectrales característicos de la muestra. Dado que la naturaleza orgánica del gel utilizado puede alterar las señales debidas a C, H, N, O, los intervalos de longitudes de onda solo tienen en cuenta las señales de componentes inorgánicos de vino, tales como líneas de Mg, Ca, K, Na. Los intervalos utilizados se muestran en la Tabla 5.2.2, en dichos intervalos no se detectaron líneas apreciables del gel.

Tabla 5.2.2: Intervalos de longitudes de onda seleccionados.

Elemento	Longitud de onda
Mg	271–291
Ca	311–340, 390–429, 513–524
Na	581–598
H _{α}	640–675
K	760–783

Modelo NN utilizado

Para llevar a cabo la clasificación se usó una red *perceptrón* multicapa *feedforward* supervisada, similar a la utilizada en el estudio de las muestras de aceite.²⁰²

La principal ventaja de este modelo de NN es la capacidad para modelar características no lineales presentes en los espectros de LIBS, así como su fácil implementación. En nuestro caso la arquitectura de NN usada consta de tres capas diferentes (de entrada, ocultas y de salida). Esta arquitectura se ha utilizado extensamente con un nivel similar de complejidad.^{203, 69, 204} El modelo de clasificación se estima mediante la calibración con un conjunto de muestras de referencia (supervisado) obteniendo los pesos y sesgos (bias) que mejor modelan todas las clases. Una vez realizada la calibración, el modelo



de NN se validó con un conjunto de datos de test (muestras no incluidas en la calibración) proporcionando el resultado de clasificación. En este estudio, se han realizado tres tipos de validación, es importante destacar que sólo un único modelo de NN ha sido desarrollado. Para su desarrollo se utilizaron el conjunto de datos de entrenamiento de cada una de las 5 clases de referencia, 80 espectros pertenecientes a la réplica 1 de las muestras M1, D1, R1, V1 y VM1. En primer lugar para evaluar el modelo se realizó una validación interna o test de sensibilidad. Para ello sólo se introdujeron como test las tres réplicas de las muestras M1, D1, R1, V1 y VM1, para la réplica 1 se usaron sólo los 20 espectros no utilizados en el entrenamiento, de modo que todos los espectros de test son desconocidos para el modelo. En segundo lugar, se evaluó la capacidad de generalización del modelo, en otras palabras, la capacidad del modelo en clasificar otros vinos pertenecientes a la misma denominación de origen pero que no han sido utilizados para el entrenamiento. Para realizarlo, se introdujeron en el modelo las librerías espectrales de 4 vinos de cada denominación de origen diferentes a los vinos de referencia (4 x 100 espectros de cada denominación de origen). Este test se considera una validación externa del modelo. Por último, se realizó un test de robustez para estudiar la capacidad del modelo a clasificar como desconocidas muestras no pertenecientes a ninguna denominación de origen usadas para su estimación. Para ello se utilizaron las librerías espectrales de vinos pertenecientes a otras denominaciones de origen españolas (CR, RB, RG, NV, SM, TR), 4 vinos de mesa y 3 vinos extranjeros (CH, DR, VC). No existe ninguna relación entre las muestras de entrenamiento con las utilizadas para este test siendo por tanto completamente independientes al modelo.

5.2.3 Resultados

Similar al caso de los aceites, se usó la correlación espectral (SC) como valor para evaluar el éxito de la clasificación. Como criterio de clasificación, consideramos una muestra correctamente clasificada cuando se obtiene un valor igual o superior al 90 % para una clase y menos al 20 % respecto a otras en términos de SC. Por el contrario, si dicho valor es inferior al 90 % la muestra se asigna como desconocida para el modelo, es decir, no perteneciente a ninguna de las *DOP* utilizadas para la estimación del modelo.

En la [Fig. 5.2.3](#) se muestran simultáneamente los espectros típicos de cada una de las 5 *DOP* utilizadas como referencia para la estimación del modelo. Se pueden apreciar ligeras diferencias entre ellos, sin posibilidad de realizar una discriminación visual de espectros.

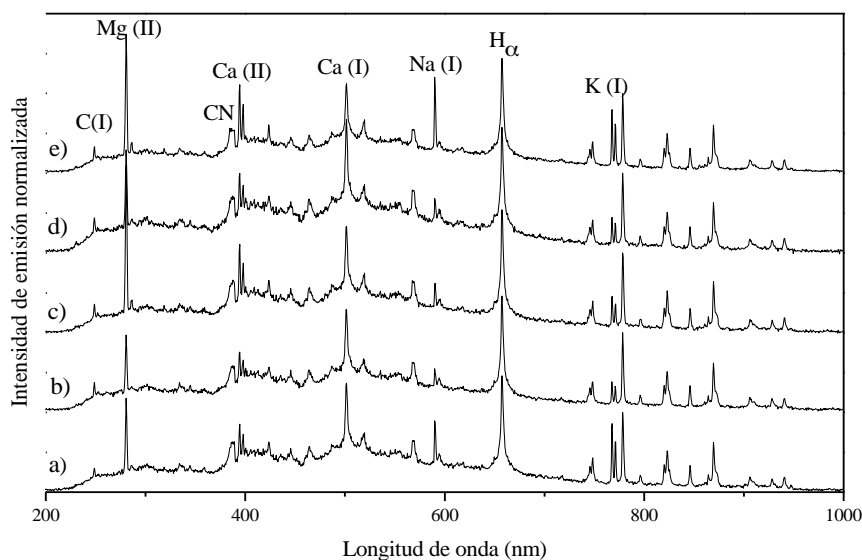


Figura 5.2.3: Espectro LIBS característico de cada Denominación de Origen utilizadas como referencia (a: Castilla, b: Ribera del Duero, c: Rioja, d: Valdepeñas, e: Vinos de Madrid)

Test de sensibilidad

La Tabla 5.2.3 muestra los resultados del análisis de sensibilidad para las tres réplicas de las muestras M1, D1, R1, V1 y VM1. El modelo es considerado sensible si presenta una tasa de clasificación correcta alta y una correlación espectral baja de espectros mal clasificados. En vista a los resultados podemos observar que el modelo presenta una alta sensibilidad, todas las réplicas son clasificadas correctamente con una correlación espectral mayor al 99 % en la clasificación e insignificante (menor al 2 %) para espectros mal clasificados. Por lo tanto, el modelo es capaz de clasificar nuevos espectros pertenecientes al mismo vino utilizado en el entrenamiento a pesar de la posible variabilidad entre los diferentes geles.

Tabla 5.2.3: Test de sensibilidad

Muestra analizada	Predicción de NN (Correlación espectral, %)					
	La Mancha	Ribera del Duero	Rioja	Valdepeñas	Vinos de Madrid	No clasificado
M1-Réplica 1	100	0	0	0	0	0
M1-Réplica 2	99	0	1	0	0	0
M1-Réplica3	99	1	0	0	0	0
D1-Réplica 1	0	100	0	0	0	0
D1-Réplica2	0	98	0	0	0	0
D1-Réplica3	0	100	0	0	0	0
R1-Réplica 1	0	0	100	0	0	0
R1-Réplica2	0	1	99	0	0	0
R1-Réplica3	0	1	98	0	1	0
V1-Réplica 1	0	0	0	100	0	0
V1-Réplica2	0	1	0	99	0	0
V1-Réplica3	1	1	0	98	0	0
VM1-Réplica 1	0	0	0	0	100	0
VM1-Réplica2	0	0	0	0	99	1
VM3-Réplica3	1	0	0	0	99	0

**Test de generalización**

Una característica común a todos métodos de clasificación multivariante, es el riesgo de generar un modelo sobreentrenado. Cuando esto ocurre, el modelo es capaz de clasificar sólo los espectros con los que ha sido entrenado, siendo incapaz de clasificar espectros de la misma muestra, aún cuando estos se han obtenido en las mismas condiciones que las referencias. Es necesario tener en cuenta esta posible limitación de los métodos supervisados. Para evitarlo en la medida de lo posible, se realizó la validación externa del modelo. La validación del modelo considerando solamente las réplicas no es suficiente, ya que los espectros no son considerados observaciones independientes. Por lo tanto, es necesario el uso de una validación externa que permita evaluar la capacidad de generalización del modelo. Para ello, se introdujeron en el modelo cuatro vinos diferentes de cada DOP. Con ello se busca una alta tasa de muestras correctamente clasificadas y una baja tasa de espectros mal clasificados o sin clasificar. La Tabla 5.2.4 muestra los resultados del test de generalización. Todos los vinos fueron correctamente clasificados con su DOP con una correlación espectral mayor al 96 %. Ningún vino fue clasificado incorrectamente o como no clasificado. Los resultados demuestran una alta capacidad de generalización y confirman que el modelo no ha sido sobre entrenado.

Tabla 5.2.4: Test de generalización

Muestra analizada	Predicción de NN (Correlación espectral, %)					
	La Mancha	Ribera del Duero	Rioja	Valdepeñas	Vinos de Madrid	No clasificado
M2	98	0	1	1	0	0
M3	100	0	0	0	0	0
M4	98	1	0	1	0	0
M5	99	1	0	0	0	0
D2	2	98	0	0	0	0
D3	0	99	1	0	0	0
D4	1	98	1	0	0	0
D5	0	100	0	0	0	0
R2	0	0	100	0	0	0
R3	1	0	98	0	0	1
R4	1	1	97	0	1	0
R5	0	2	97	1	0	0
V2	0	0	0	100	0	0
V3	2	1	0	97	0	0
V4	1	0	0	98	0	1
V5	0	0	0	100	0	0
VM2	1	0	0	1	98	0
VM3	0	0	0	0	99	1
VM4	0	0	1	1	98	0
VM5	0	0	0	0	100	0

Test de robustez

Para evaluar la robustez del modelo de NN, se midieron un total de trece muestras de vino, de las cuales 6 pertenecen a DOP diferentes a las usadas en el entrenamiento, 4 vinos de mesa y tres vinos extranjeros. Con este test se evaluó la respuesta del modelo para detectar muestras externas completamente independientes, como desconocidas. La Tabla 5.2.5 presenta los resultados del test de robustez del modelo. En primer lugar cabe destacar que ningún vino ha sido clasificado como perteneciente a una DOP usada en la estimación del modelo a pesar de su proximidad geográfica, como por ejemplo Rioja y vinos de Navarra, siendo todos los vinos clasificados correctamente como desconocido. En segundo lugar, los vinos de mesas, generalmente elaborados con mezclas de uva de diferente procedencia pudiendo en

casos excepcionales contener uva con DOP, ninguno ha sido clasificado incorrectamente como perteneciente a alguna DOP de referencia y todos ellos son identificados como desconocidos. Sólo en caso de TW3 se obtuvo una correlación espectral de 70 % a la clase La Mancha debido a la mezcla de uvas, en la mayoría de los casos, de la región de La Mancha. Sin embargo, el SC es inferior al umbral fijado y por lo tanto clasificado como desconocido para el modelo. Por último, ningún vino extranjero ha sido clasificado y todos son considerados desconocidos para el modelo. A la vista de estos resultados podemos decir que el modelo de NN presenta una alta robustez cuando es sometida a la predicción de muestras externas.

Tabla 5.2.5: Test de Robustez

Muestra analizada	Predicción de NN (Correlación espectral, %)					Correctamente como desconocido
	La Mancha	Ribera del Duero	Rioja	Valdepeñas	Vinos de Madrid	
CR	5	0	5	0	0	✓
RB	0	0	8	0	0	✓
RG	4	0	2	0	0	✓
NV	2	0	0	2	0	✓
SM	7	3	0	0	5	✓
TR	40	0	2	0	3	✓
TW1	32	0	0	0	0	✓
TW2	12	0	32	0	0	✓
TW3	70	0	0	0	10	✓
TW4	22	0	0	0	8	✓
CH	0	0	0	7	0	✓
DR	0	8	0	2	0	✓
VC	0	0	0	7	0	✓

5.2.3 Principales conclusiones de la identificación de vinos

La combinación LIBS/NN proporciona una metodología adecuada, sencilla y económica de alta sensibilidad y precisión en la discriminación de las denominaciones de origen y/o detección de fraudes en vinos. El método desarrollado tiene aplicación en diversos campos relacionados con la industria vitivinícola como el control de calidad y



la seguridad alimentaria. El método desarrollado permite la identificación instantánea de muestras de vino basadas en los espectros LIBS obtenidos y considerados como “huella digital espectral” de la muestra. El proceso de gelificación produce mejores resultados que el análisis de las muestras líquidas directamente y es la primera vez que se describe como alternativa para analizar y mejorar las condiciones de trabajo de la técnica LIBS.

La reducción del número de variables es un método adecuado para dotar a la red neuronal de una forma de ajustar su nivel de discriminación. El uso del espectro completo permite una mayor capacidad de discriminación en detrimento de una menor capaz de generalizar. El número de variables y la selección de las mismas afecta de forma significativa a los resultados obtenidos, su selección es de vital importancia para dotar a la red tanto de una alta capacidad de clasificación como de una habilidad de generalizar adecuada. Las variables seleccionadas por tanto contienen información de las condiciones climáticas y la ubicación geográfica de los viñedos, del sistema bioquímico propio de la vid y las particularidades del cultivo, del suelo y las reacciones enzimáticas asociadas con efectos específicos cinéticos y termodinámicos en diversos pasos del metabolismo durante la fermentación y almacenaje del vino.

Se han realizados diferentes validaciones al modelo de NN generado, siendo adecuado para la discriminación inequívoca de vinos con denominaciones de origen protegidas.

El método de análisis de vinos para su discriminación e identificación por ablación láser puede ser también aplicado a otras bebidas alcohólicas como vinos rosados, blancos, espumosos, whisky, ron, vodka, ginebra o tequilas.



Capítulo 6:

LIBS aplicado a la ciencia forense

Según su definición, la ciencia forense es la aplicación de prácticas científicas dentro del proceso legal. A través de la ciencia forense, se identifican sospechosos y se desvelan las circunstancias que rodean un crimen aplicando conocimientos científicos abarcando varias áreas de conocimiento y la colaboración de científicos e investigadores altamente cualificados. Entre las distintas disciplinas que engloba la ciencia forense se puede destacar algunas como la toxicología forense, que se ocupa de la presencia o ausencia de drogas y químicos en el organismo de un individuo, la contabilidad forense, que investiga problemas relacionados con fraudes, la fauna forense, capaz de resolver crímenes contra animales, el arte forense, donde se aplica el arte para buscar criminales o personas desaparecidas y en la identificación de víctimas, la ingeniería forense, que aplica conceptos ingenieriles a la resolución de circunstancias legales, la patología forense, ocupada del estudio de enfermedades relacionadas con la muerte de un individuo a través de autopsias para determinar la causa y forma de muerte, y por último, la antropología forense, siendo “la rama de la antropología física que con fines forenses trata de la identificación de restos más o menos esqueléticos, humanos o de posible pertenencia humana”.²⁰⁵

En cuanto a la identificación de personas fallecidas, no suele plantear grandes dificultades cuando se encuentran cadáveres aislados y de fallecimiento reciente, pero puede resultar complicada, problemática y llena de dificultades cuando hay que identificar múltiples víctimas en catástrofes en masa o atentados terroristas.



En España la Ley 2/85, de Protección Civil, el Real Decreto 1378/85 y las Normas Básicas de Protección Civil de 24-10-91 no contiene ni una sola referencia expresa a la identificación de las víctimas tras un suceso catastrófico. Tampoco se refiere para nada a los procedimientos técnicos empleados para la identificación de las víctimas. La Organización Internacional de Policía Criminal, más conocida como INTERPOL, en su "Manual de Identificación de Víctimas de Catástrofes"²⁰⁶ recomienda la formación y creación de comisiones permanentes para identificación de víctimas de catástrofes que permitan la planificación anticipada de las acciones que hayan de emprenderse, el entrenamiento de grupos especiales, así como los procedimientos de análisis necesarios para llevarlo a cabo.

Los progresos de las diversas técnicas empleadas en la identificación de cadáveres permiten dar soluciones, con mayor rapidez y grado de fiabilidad, a la problemática concreta de cada suceso puntual. La identificación de víctimas implica la participación de personal de diversas áreas como la medicina legal y forense, odontología forense, radiología, antropología, biología, química, física, etc, siendo, por tanto, altamente multidisciplinar.

Entre los métodos y técnicas más aplicadas en la actualidad para la identificación de víctimas en masa se pueden destacar:²⁰⁶

- Reconocimiento visual (cuando sea posible)
- Reconocimiento de objetos y prendas
- Descripción física, particularidades y rasgos específicos del individuo
- Identificación por huella digital (lofoscópica)
- Odontología forense
- Radiológica



- Serológicas y técnicas analíticas

Algunos de estos métodos, como el reconocimiento visual por parte de un familiar, no son definitivos, sólo las huellas dactilares, la radiografía bucal post mortem, los números de serie de prótesis implantadas o marcapasos y los análisis de ADN son considerados definitivos. Los análisis de ADN permiten identificar a los cadáveres más deteriorados e incluso recomponer aquellos cuyos miembros quedaron esparcidos. Sin embargo, en múltiples ocasiones no es posible obtener el material genético necesario para llevar a cabo el análisis, impidiendo la identificación por ADN.²⁰⁷ Los principales inconvenientes encontrados a la hora de llevar a cabo los análisis genéticos son la falta de la cantidad de ADN necesaria, entre 20 y 100 ng, sobre todo cuando los restos biológicos encontrados son de pequeño tamaño o están degradados. Por otra parte el tiempo mínimo requerido para el análisis es en general de dos a tres días para poder emitir un resultado confiable. Es importante destacar también como las etapas de preparación de las muestras hacen que la metodología presente un mayor uso de recursos temporales y económicos.

Es por ello que se ha propuesto la utilización de la técnica LIBS para el desarrollo de procedimientos analíticos capaces de realizar el análisis forense disminuyendo los costes tanto temporales como económicos. A continuación, se detallan los procedimientos empleados en esta tesis para llevar a cabo dicho análisis.

Los resultados comentados han sido publicados con en *Spectrochimica Acta Part B Atomic Spectroscopy* con la referencia **101** (2014) 21-25 y en *Chemometric and Intelligent Laboratory Systems* **146** (2015) 354-364.

6. 1 Discriminación de cuerpos humanos en base a huesos y dientes por espectroscopia de ablación láser (LIBS) y Redes Neuronales (NN)

Resumen

Se ha desarrollado y aplicado un método rápido y mínimamente destructivo basado en la espectroscopia de ablación inducida por láser (LIBS) y redes neuronales (NN) para la clasificación y la discriminación de huesos y fragmentos de dientes humanos. Esta metodología es de gran utilidad en tareas de identificación de víctimas en catástrofes (DVI). Se ha observado que la composición elemental de las muestras de huesos y dientes proporciona suficiente información para obtener una correcta discriminación y reagrupamiento de los restos. Todos los individuos fueron clasificados con una correlación espectral superior al 95%, independientemente del tipo de hueso o diente analizado. No se observaron falsos positivos ni tampoco falsos negativos, lo que demuestra la alta robustez y precisión de la metodología propuesta.

6.1.1 Introducción a la Identificación de restos óseos mediante LIBS y NN

La identificación inequívoca de cuerpos humanos es de gran importancia desde el punto de vista tanto legal como administrativo. En ocasiones, cuando el número de restos encontrados es grande (desastres naturales o genocidios) dicha identificación es extremadamente difícil.^{207,208} En situaciones tales como accidentes aéreos, desastres naturales o fosas comunes es de vital importancia “poner nombre” a cada víctima encontrada, no solo desde el punto de vista humanitario sino también legal para la resolución de litigios. Para ello, la identificación de individuos sigue las metodologías y protocolos descritos en tratados como la guía de Identificación de víctimas en desastres de la Interpol.²⁰⁶ El proceso de identificación de víctimas es altamente multidisciplinar, haciendo uso de diferentes grupos de expertos y diversas técnicas como el análisis de



las crestas papilares, patología forense, odontología forense y técnicas genéticas, entre otras. En situaciones donde los cuerpos se encuentran en buen estado de conservación, la identificación se realiza en base a marcas, tatuajes, prendas y joyas que porta la víctima, así como la identificación visual por parte de un familiar y a través de registros médicos.²⁰⁸⁻²¹⁰ Sin embargo, en casos donde los restos presentan un alto grado de putrefacción, degradación o cualquier tipo de destrucción, los métodos habituales no son adecuados para llevar a cabo la identificación y es necesario el uso de metodologías de análisis más avanzadas. En el 42% de los casos, los restos encontrados no permiten llevar a cabo la identificación o bien no se cuenta con registros médicos pre-mortem y es necesario el uso de técnicas genéticas.^{211, 212} La imposibilidad de obtener DNA de parientes relacionados con las víctimas, la complejidad de la obtención del material genéticos de restos óseos y dentales así como su degradación y conservación pueden dificultar el proceso de identificación a través de técnicas genéticas.²¹³ Además, cuando el número de muestras a analizar es elevado, la identificación por DNA se vuelve tediosa y difícil, involucrando largos tiempos de análisis y siendo económicamente costosa.

Por todo lo expuesto, existe la necesidad de una metodología rápida, simple, directa y económica que permita agilizar las labores de agrupación e identificación de restos humanos. La técnica LIBS es capaz de obtener un huella espectral de la muestra con un único pulso de forma instantánea permitiendo junto con el uso de herramientas de procesamiento de datos la clasificación de muestras, suscitando un enorme interés en la ciencia forense.^{73, 186, 214-215}

Los restos óseos y dentales, debido a su alta resistencia y su baja degradación se encuentran frecuentemente en escenarios forenses, siendo evidencias útiles para la



antropología, arqueología y las ciencias forenses en general.^{66, 23} Aunque las diferencias en la composición elemental de los restos de diferentes individuos son mínimas, algunos elementos traza están relacionados con la dieta y el entorno del individuo,²¹³ produciendo ligeras variaciones en el espectro de emisión y permitiendo su discriminación a través de un análisis quimiométrico adecuado. Las redes neuronales han demostrado ser una buena solución para el desarrollo de aplicaciones en diversas áreas, gracias a su capacidad de generalizar y su relativamente sencilla implementación.^{13,190-191}

El objetivo de este trabajo es el desarrollo de un método de análisis simple, directo y económico basado en LIBS y NN para discriminar y reagrupar individuos usando restos óseos y dentales.

6.1.2 Procedimiento Experimental

El sistema experimental utilizado en este estudio es el mismo que el descrito con anterioridad en casos similares como por ejemplo el análisis de aceites.²⁰² Es este caso para prevenir el efecto del *bremsstrahlung*, se fijó el tiempo de retardo entre el pulso láser y la adquisición del espectro en 4 μ s, dicho valor fue optimizado en función de la relación señal ruido a diferentes retrasos.

Muestras óseas y dentales

Para este estudio, se analizaron los restos de 9 individuos, considerando veinticinco fragmentos hueso de cinco individuos y doce dientes de cuatro individuos diferentes. Todos los restos fueron recogidos en el osario del cementerio de Segovia con el permiso de las autoridades locales. Los restos de tejido y suelo de las muestras se eliminaron con un cepillo fino y posteriormente se lavaron con agua templada sin usar ningún tipo de jabón o producto químicos que pudiera interferir en la composición.



Después, se enjuagaron varias veces con agua destilada y se dejaron secar al aire durante un día antes de su medida. La Tabla 6.1.1 muestra la nomenclatura y tipo de hueso y diente usados en el estudio. La nomenclatura empleada utiliza dos números y una letra: la letra (B o T) indica el tipo de muestra, hueso (B) y diente (T). El primer número indica el individuo de pertenencia y el segundo el hueso o diente del individuo considerado.

Medidas LIBS y librerías espectrales

Todos los restos fueron medidos a presión atmosférica y temperatura ambiente. Cada espectro LIBS fue obtenido tras someter la muestra a 6 pulsos en la misma posición. Los tres primeros pulsos se utilizaron como etapa de limpieza y eliminación de posibles restos sobre la superficie y los tres siguientes fueron promediados para su adquisición. Posteriormente la muestra se desplazó 0.25mm en el eje X-Y para irradiar una nueva zona, evitando la interferencia con zonas previamente analizadas. Para cada muestra se obtuvo un conjunto de 50 espectros LIBS. Solo en el caso del fémur derecho y el diente incisivo de todos los individuos, se obtuvieron dos conjuntos de 50 espectros. El primero de ellos se utilizó para el entrenamiento del modelo de NN, mientras que el segundo se utilizó para la validación de los modelos. Todos los espectros fueron normalizados con la intensidad de la línea de Ca a 393.37 nm para evitar las variaciones espectrales producidas por fluctuaciones en la energía del láser.

**Tabla 7.1.1:** Descripción de las muestras

Individuo	Identificación Huesos	Descripción
1	11B	Fémur derecho
	12B	Segunda Falange
	13B	Primer metatarso
	14B	Cúbito
	15B	Costilla derecha
2	21B	Fémur derecho
	22B	Radio
	23B	Escápula
	24B	Primera falange
	25B	Cúbito
3	31B	Fémur derecho
	32B	Cúbito
	33B	Clavícula izquierda
	34B	Primera falange
	35B	Húmero derecho
4	41B	Fémur derecho
	42B	Cúbito derecho
	43B	Tibia derecha
	44B	Fémur izquierdo
	45B	Radio izquierdo
5	51B	Fémur derecho
	52B	Calcáneo
	53B	Cúbito derecho
	54B	Metacarpo
	55B	Clavícula izquierda

Individuo	Identificación Dientes	Descripción
6	61T	Incisivo
	62T	Canino
	63T	Premolar
7	71T	Incisivo
	72T	Canino
	73T	Premolar
8	81T	Incisivo
	82T	Canino
	83T	Premolar
9	91T	Incisivo
	92T	Canino
	93T	Premolar

Modelo NN utilizado

Se desarrolló un modelo de red neuronal propia (anexo 2) en Matlab (Mathworks, 2010a) capaz de clasificar muestras de huesos y dientes con su individuo de pertenencia a través del análisis de los espectros LIBS. La arquitectura de red utilizada se basa en una red multicapa tipo *perceptrón*, con conexiones hacia adelante con un algoritmo de entrenamiento supervisado. La topología de la red consiste en tres capas de neuronas (entrada, oculta y salida), donde cada neurona está conectada con todas las neuronas de la capa anterior. Las conexiones neuronales son modificadas a través de los pesos sinápticos y sesgos modulando su respuesta y transfiriendo la información a las neuronas de la siguiente capa a través de una función de transferencia de tipo sigmoideal. El algoritmo de entrenamiento utilizado fue *back-propagation* (BP) usando el método de gradiente conjugado,^{13, 192} haciendo uso de las segundas derivadas para minimizar la función de error de las variables. Para detener el entrenamiento, se usó el error cuadrático medio (MSE) entre los datos de entrada y la salida de la red neuronal evitando así el sobreajuste de la NN.¹⁹³

Entrenamiento de la NN

Debido a que la red utilizada neuronal empleada utiliza un entrenamiento supervisado, es necesario el uso de espectros de referencia representativos de la muestras para la estimación del modelo. Se realizaron dos modelos independientes, una para el caso de huesos y otro para dientes. En el caso de las muestras óseas, se usaron las cinco librerías espectrales de referencia pertenecientes a cada fémur derecho de cada individuo (50 espectros x 5 individuos) y una matriz de clases que contiene información sobre la clase de pertenencia de cada espectro de referencia. La dimensión de esta matriz, tiene el mismo número de filas que espectros contiene la matriz de

entrenamiento y columnas similar al número de clases. En el caso de los dientes, se usó la librería espectral de entrenamiento del diente incisivo de cada individuo (50 espectros x 4 individuos) y su correspondiente matriz de referencia. Como parte del proceso interno del entrenamiento, el conjunto de datos se divide en un dos conjuntos, el 80% se utilizó para el entrenamiento y el 20% como validación propia del proceso de entrenamiento. Una vez generados ambos modelos, se validaron usando las librerías espectrales del conjunto de espectros no usados en el entrenamiento del modelos y del resto de muestras.

Validación del modelo de NN

La validación de los modelos de NN se realizó con el objetivo de evaluar la calidad de los modelos y su capacidad para predecir la pertenencia de muestras no introducidas previamente en el entrenamiento, así como la discriminación de restos óseos y dentales en función a su individuo de pertenencia.

Los modelos se evaluaron en función de su exactitud o tasa de discriminación positiva, definiendo la correlación espectral como el número de espectros clasificados correctamente con su clase sobre el total de espectros de la muestra.¹⁹⁰ Para considerar un muestra clasificada correctamente, se fijó arbitrariamente un límite mínimo del 90 % en el valor de la correlación espectral.

6.1.3 Resultados

Espectro de hueso y diente

La [Fig. 6.1.1 \(a\) y \(b\)](#) muestra un espectro típico normalizado de una muestra de hueso y diente, respectivamente. Las líneas de emisión de los principales elementos a sus correspondientes longitudes de onda fueron identificadas en base a las referencia de NIST.⁶⁸ Se identificaron líneas correspondientes a C I (247.85), Mg I (285.21; 518.36),



Mg II (280.27), Ca I (422.67; 445.48; 551.30; 560.13), Ca II (393.37; 396.85), Sr I (407.77), Na I (588.99; 589.59), O I (777.42; 844.62) y N I (746.83; 821.63; 868.34). En ambos casos, la línea más intensa corresponde a la emisión de calcio debido a la composición de la muestra, principalmente hidroxiapatita ($\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2$) y fosfato de calcio.^{216,217} Las cantidades relativas de los diferentes elementos puede variar, y de hecho varía, entre los diferentes individuos e incluso dentro de un mismo individuo debido a la diferente procedencia embriológica y metabolismo de algunos tipos de hueso.²¹³ Por otra parte, debido a que los espectros fueron tomados a presión atmosférica, se pueden observar líneas de emisión debidas a oxígeno y nitrógeno, presente alrededor del plasma. Se ha demostrado que dichas líneas atmosféricas no producen interferencias espectrales con las intensidades relativas de las líneas del resto de elementos presentes en las muestras.¹⁹⁷ También es observable un continuo de radiación, que es un componente típico de todos los espectros LIBS.

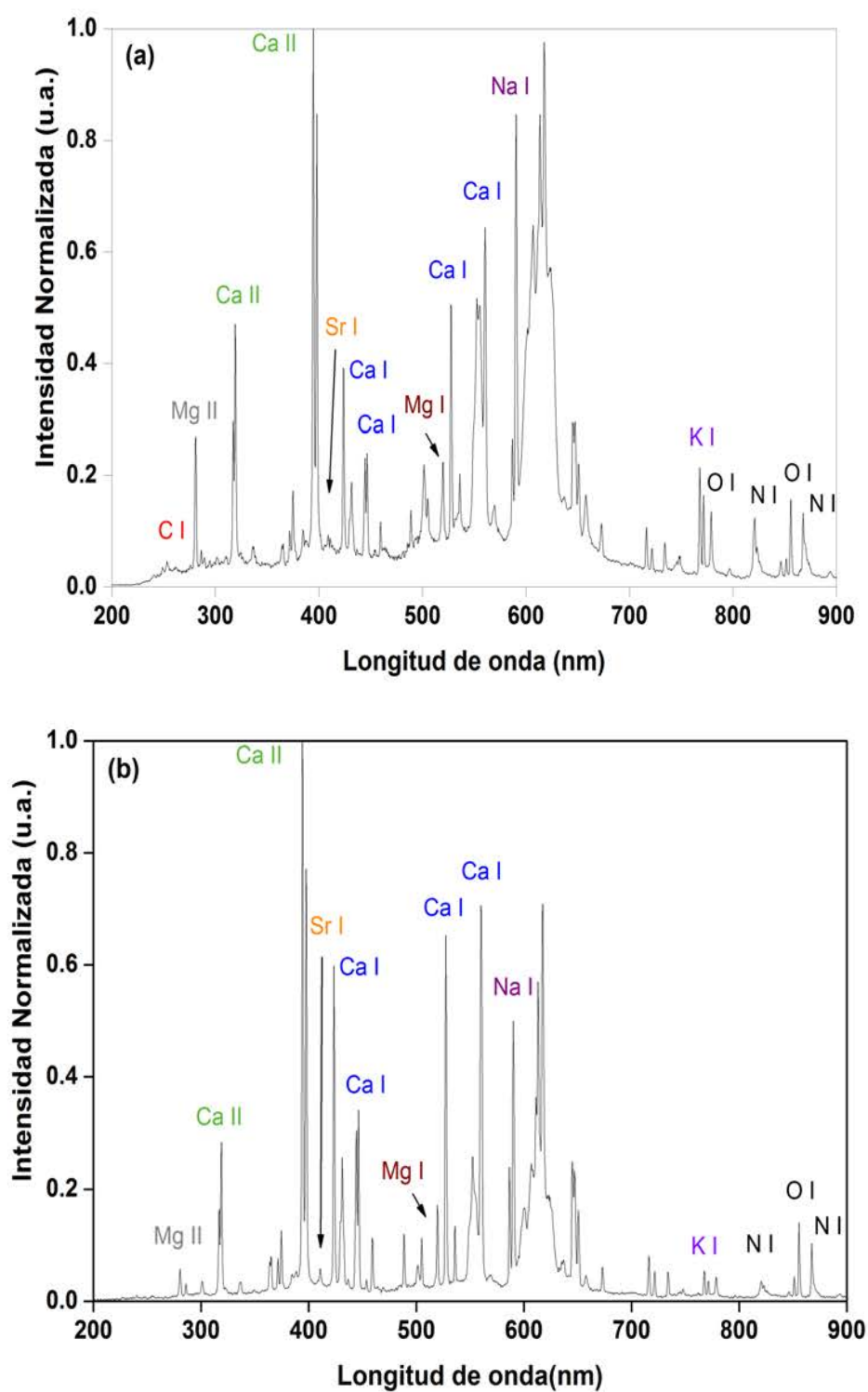


Figura 6.1.1: Espectro típico normalizado LIBS de una muestra de a) hueso y b) diente.

Discriminación de individuos

Como variables de entrada, intensidades a cada longitud de onda, se seleccionaron tres intervalos de longitud de onda englobando las principales líneas de emisión correspondientes a Ca, Sr y Mg. Los intervalos utilizados fueron: 390-410nm, 420-480nm y 516-532nm. Se utilizó una reducción de variables con el objetivo de eliminar datos redundantes que pueden producir una disminución en la capacidad de discriminación del modelo de NN. Un total de 195 variables fueron finalmente utilizadas para el entrenamiento de los modelos de NN, 10 neuronas en la capa oculta y en la capa de salida, el número de individuos para cada caso. Se generaron dos modelos independientes para cada tipo de muestra (huesos y dientes).

Discriminación de individuos a través de muestras óseas

El entrenamiento del modelo de NN se realizó con las librerías espectrales correspondientes a los 5 fémures derechos de los individuos del 1 al 5, generando una matriz de 50 espectros por 5 individuos (250 columnas) y 195 variables (filas). Cada individuo es asignado un número de identificación para realizar la estimación del modelo de forma supervisada. La [Fig. 6.1.2a](#), muestra el resultado de la autovalidación del modelo de NN durante el proceso de entrenamiento. A pesar de las pequeñas oscilaciones en los valores de salida de NN, todos los espectros son asignados correctamente con su clase de pertenencia con una incertidumbre menor al 1%.

Una vez el modelo de NN ha sido estimado, se realizó la predicción del resto de muestras introduciendo las librerías espectrales de test y comparando las características espectrales de los nuevos espectros.

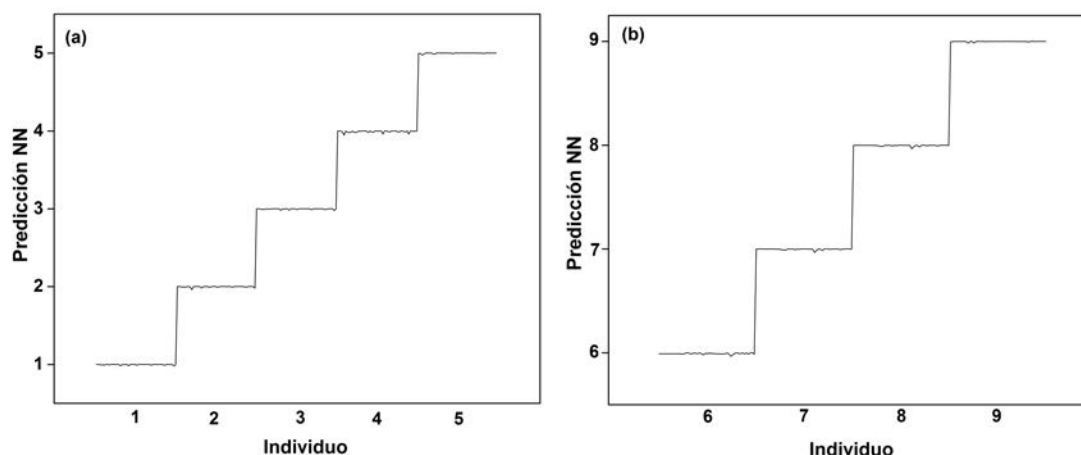


Figura 6.1.2: Resultados entrenamiento de la NN, a la izquierda (a) modelo de NN de muestras óseas y derecha (b) modelo para muestras dentales

La Tabla 6.1.3 muestra los resultados de la NN en el agrupamiento de individuos a través de restos óseos. Todas las muestras fueron clasificadas correctamente con una correlación espectral superior al 98 % demostrando la capacidad de generalización de la NN. Es importante recalcar que el modelo sólo se estimó con un único hueso por individuo y que a pesar de las diferencias espectrales inter e intra-individuos el modelo es capaz de agrupar huesos desconocidos con su individuo de pertenencia. Es importante mencionar que los espectro no clasificados correctamente (menos del 1% del total) no son asignado como pertenecientes a otro individuo, si no que quedan sin asignar y por tanto clasificados como desconocidos para el modelo. La alta capacidad para generalizar y la baja tasa de espectros mal asignados demuestran la robustez y exactitud de la metodología propuesta en el agrupamiento y discriminación de individuos a través de sus huesos.

Se realizó un test de robustez al modelo de NN. Para ello se suprimió la librería espectral de un individuo y una vez generado un modelo nuevo se introdujo este para su predicción, repitiendo este proceso de forma iterativa para cada individuo. Mediante



este test se evalúa la capacidad del modelo a detectar muestras desconocidas así como asegurar que no se produzcan falso positivos (clasificado como perteneciente a un individuo sin serlo). En todos los casos las muestras eliminadas del entrenamiento fueron clasificadas como desconocidas.

Los resultados mostrados en la Tabla 6.1.2 demuestran que a pesar de existir diferencias elementales en la composición de los huesos dentro de un individuo, las diferencias composicionales entre individuos es significativa y permite la realización del agrupamiento de huesos basado en las medidas LIBS de sus huesos.

A diferencia de otros métodos,²¹³ una importante ventaja de esta metodología es la posibilidad de discriminar entre individuos sin la necesidad de una separación previa de las muestras en función del tipo de hueso. Además, el estado de conservación de los restos no afecta al sistema LIBS/NN, como en la mayoría de casos cuando se usa DNA.^{218- 220}

**Tabla 6.1.2:** Resultado predicción NN para las muestras de huesos

Muestra analizada	Predicción de NN (Correlación espectral, %)					
	Ind.1	Ind.2	Ind.3	Ind.4	Ind.5	No asignado
11B	100	0	0	0	0	0
12B	100	0	0	0	0	0
13B	100	0	0	0	0	0
14B	100	0	0	0	0	0
15B	100	0	0	0	0	0
21B	0	100	0	0	0	0
22B	0	98	0	0	0	2
23B	0	100	0	0	0	0
24B	0	100	0	0	0	0
25B	0	98	0	0	0	2
31B	0	0	98	0	0	2
32B	0	0	100	0	0	0
33B	0	0	98	0	0	2
34B	0	0	100	0	0	0
35B	0	0	100	0	0	0
41B	0	0	0	100	0	0
42B	0	0	0	98	0	2
43B	0	0	0	100	0	0
44B	0	0	0	98	0	2
45B	0	0	0	100	0	0
51B	0	0	0	0	100	0
52B	0	0	0	0	86	14
53B	0	0	0	0	98	2
54B	0	0	0	0	100	0
55B	0	0	0	0	98	2

Discriminación de individuos a través de muestras dentales

El entrenamiento del modelo de NN se realizó con las librerías espectrales correspondientes al diente incisivo de cada uno de los cuatro individuos del 6 al 9, generando una matriz de 50 espectros por 4 individuos (200 columnas) y 195 variables (filas). Similar al caso de muestras óseas, cada espectro es asignado a un número de identificación para cada individuo. La [Fig. 6.1.2b](#) muestra el resultado de la predicción del modelo de NN durante el proceso de entrenamiento. Como en el caso de huesos, el

En general, los resultados obtenidos demuestran que la combinación LIBS/NN es una metodología rápida, simple y económica para la discriminación y agrupamiento de individuos a través de restos óseos y dentales. No produciendo falsos positivos ni falsos negativos demostrando una alta robustez.

Muestra analizada	Predicción de NN (Correlación espectral, %)				
	Ind.1	Ind.2	Ind.3	Ind.4	No asignado
61T	100	0	0	0	0
62T	100	0	0	0	0
63T	100	0	0	0	0
71T	0	98	0	0	2
72T	0	100	0	0	0
73T	0	96	0	0	4
81T	0	0	100	0	0
82T	0	0	94	0	6
83T	0	0	100	0	0
91T	0	0	0	100	0
92T	0	0	0	100	0
93T	0	0	0	100	0



6.1.4 Conclusiones a la discriminación de individuos mediante LIBS y NN

Se ha desarrollado un método basado en LIBS combinado con redes neuronales supervisadas para la discriminación rápida y robusta de restos biológicos humanos en base a las características espectrales de huesos y dientes.

La técnica no pretende competir con las metodologías de análisis actuales, generalmente DNA, la combinación LIBS/NN podría ser considerada como una técnica complementaria sencilla y rentable que ayude a los métodos biológicos, más lentos y caros, en la discriminación de restos humanos. El método permite llevar a cabo una discriminación post-mortem, ya sea utilizando restos de hueso o diente, gracias a la huella espectral de cada individuo generada por LIBS, sin embargo esta metodología no permite relacionar los restos con los parientes vivos siendo necesaria la aplicación de técnicas genéticas. A pesar de la complejidad de la matriz y la cantidad de líneas presentes en el espectro LIBS, la selección única de líneas correspondientes a calcio, magnesio y estroncio permiten la discriminación entre individuos con una alta precisión.

Gracias a su velocidad, robustez, carácter mínimamente destructivo, micro-análisis y facilidad de uso, la combinación de LIBS con métodos quimiométricos de análisis ofrece ventajas significativas en el ámbito de la ciencia forense, facilitando especialmente las tareas de reconstrucción y discriminación de individuos víctimas de desastres naturales y actos de genocidio.

Resumen

Se ha realizado una comparativa entre diferentes métodos quimiométricos para la discriminación de restos óseos humanos basada en medidas espectrales obtenidas a través la técnica LIBS. Se ha estudiado la combinación entre LIBS y siete métodos de clasificación supervisados en una aplicación real como la discriminación de muestras óseas con composición elemental similar, obtenidas en un osario y bajo los efectos de parámetros no controlados como precipitaciones, cambios de temperatura y acción de insectos, entre otros. Se ha estudiado un total de 25 restos óseos pertenecientes a 5 individuos. Cada método de clasificación ha sido evaluado mediante tres procesos de validación, teniendo en cuenta la sensibilidad, la capacidad de generalización y la robustez. La exactitud de cada método fue en orden creciente: 42% para Análisis Discriminante Lineal (LDA), 56% para *Support Vector Machines* (SVM), 58% para *Soft Independent Modeling of Class Analogy* (SIMCA), 58% para *Partial least Squares–Discriminant Analysis* (PLS-DA) y 100% para Redes Neuronales Artificiales (NN). Los resultados obtenidos muestran que la Red Neuronal supera al resto de métodos en términos de sensibilidad, capacidad de generalización y robustez, mientras que SIMCA, PLS-DA, LDA y SVM no presentan suficiente exactitud para llevar a cabo la discriminación entre restos óseos con un alto grado de similitud.



6.2.1 Introducción a la comparativa de métodos quimiométricos para la discriminación y agrupamiento de restos óseos.

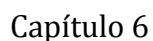
La quimiometría se ha establecido como una metodología madura para extraer información significativa y útil de enormes conjuntos de datos. La quimiometría en química engloba multitud de enfoques como la inferencia estadística, calibración univariante y multivariante, diseño experimental, procesamiento de señales, optimización secuencial y simultánea y procesos de clasificación, ya sean supervisados o no supervisados.^{100-102, 221} Con el objetivo de llevar a cabo la mejor clasificación posible a partir de los datos originales, existen multitud de algoritmos estadísticos y computacionales. La selección del método de clasificación más adecuado depende de la aplicación analítica y de la información requerida. Como se comentó en el capítulo 3 los métodos de clasificación se pueden clasificar como supervisados y no supervisados. Para resumir podemos decir que los métodos de clasificación no supervisados operan con una única matriz de datos (objetos) y no se asume ningún modelo de clasificación previo, mientras que los métodos supervisados trabajan con dos matrices, una de objetos (input) y otra de clases (target). El método estima un modelo de clasificación en función de la relación entre objetos y clases, representando las características que definen el proceso de clasificación. La predicción de nuevos objetos (desconocidos para el modelo) se realiza en base a la similitud a cierta clase con la que el modelo fue estimado. En cuanto a los métodos de clasificación supervisado podemos diferenciar entre métodos lineales y no lineales. Los métodos lineales usan combinaciones lineales de las variables para estimar el modelo de clasificación; mientras que los métodos no lineales aplican algoritmos complejos incrementando la sofisticación y la capacidad de clasificación del método.

Se ha elegido la técnica LIBS debido a su capacidad de producir una huella espectral digital de la composición elemental de la muestra, la cual ha permitido obtener resultados previos en la clasificación en tiempo real de diversas muestras como bacterias^{13, 69}

Existen multitud de métodos quimiométricos y por tanto cada autor aplica aquellos que piensa que mejor funcionan o bien aquellos que son más accesibles y sencillos de implementar.^{107, 103, 222} Sin embargo, no existe ningún convenio o guía para la selección del método de clasificación más adecuado. La selección del método de clasificación no es directa y los datos obtenidos han de ser estudiados cuidadosamente con el fin de seleccionar aquel que mejor se adapte y proporcione los mejores resultados de clasificación.

Para evaluar los métodos de clasificación existen diversas técnicas y procesos. En general un modelo de clasificación adecuado ha de ser capaz de reconocer correctamente aquellas muestras con las que fue entrenado (alta sensibilidad), además la predicción de muestras de la misma clase no incluidas en el entrenamiento ha de ser correcta y por tanto con una alta capacidad de generalizar. Por otra parte, es importante la obtención de un modelo de clasificación robusto, en otras palabras, que sea capaz de reconocer como desconocidas muestras que no pertenezcan a ninguna de las clases con la que fue entrenado.

La combinación entre la técnica LIBS y diversos métodos quimiométricos ya se ha aplicado en bibliografía, sin embargo no todos los posibles métodos han sido estudiados, los métodos más comunes en la comunidad LIBS son SIMCA, LDA, PLS-DA, funciones discriminantes y en menor medida SVM y NN.^{115, 215, 223-225} Se han desarrollado todo tipo de aplicaciones incluyendo el estudio de rocas marcianas en



La sensibilidad, capacidad de generalizar y robustez de todos los modelos fue evaluada para determinar el método de clasificación más adecuado al análisis.



6.2.2 Métodos aplicados

El fundamento teórico de todos los métodos estudiados se ha descrito con anterioridad en el capítulo de métodos quimiométricos y por tanto para una mayor información de los mismos ruego se dirija a dicha sección (Capítulo 3).

6.2.3 Procedimiento experimental

Muestras Óseas

Veinticinco huesos de cinco individuos fueron recogidos del cementerio local de Segovia cumpliendo con todos los requisitos y con las autorizaciones pertinentes. El tratamiento de muestra fue similar al aplicado en el trabajo anterior (Sección 6.1). La identificación de cada muestra incluye dos números, el primero de ellos (1-5) corresponde al individuo estudiado y el segundo al tipo de hueso (1-5), se muestra en la Tabla 6.2.1.

Sistema experimental

El sistema experimental utilizado en este estudio se describió en el apartado 6.1. Como resumen se utilizó un láser de Nd:YAG con sistema de Q-switch a 1064nm y con pulso de 4 ns (FWHM). La fluencia se fijó a 20 J/cm² a través de una lente plano convexa de 10 cm de distancia focal produciendo un spot de unas 100 µm. La emisión del plasma generado es recogida a través de una fibra óptica de sílice fundida acoplada a un colimador con distancia focal de 7 mm que se colocó a 4 cm de la muestra con un ángulo de 45° sobre la normal de la superficie. El otro extremo de la fibra óptica se acopló a un espectrómetro (StellarNet, Tampa, FL, USA) con una red de 300 l/mm y que produce una resolución de 0.5 nm cuando usamos una rendija de entrada de 7 µm, permitiendo obtener una ventana espectral de 200 a 1000 nm. Para evitar la primera emisión del plasma, efecto Bremsstrahlung inverso, se produjo un retraso de 2 µs entre

la obtención del espectro y el pulso del láser. El retraso, tiempo de integración y la fluencia fueron optimizadas en función de la relación señal/ruido.

Tabla 6.2.1: Muestras óseas utilizadas en el análisis

Individuo	Identificación Huesos	Descripción
1	11	Fémur derecho
	12	Segunda Falange
	13	Primer metatarso
	14	Cúbito
	15	Costilla derecha
2	21	Fémur derecho
	22	Radio
	23	Escápula
	24	Primera falange
	25	Cúbito
3	31	Fémur derecho
	32	Cúbito
	33	Clavícula izquierda
	34	Primera falange
	35	Húmero derecho
4	41	Fémur derecho
	42	Cúbito derecho
	43	Tibia derecha
	44	Fémur izquierdo
	45	Radio izquierdo
5	51	Fémur derecho
	52	Calcáneo
	53	Cúbito derecho
	54	Metacarpo
	55	Clavícula izquierda

Medidas LIBS y librerías espectrales

Las muestras se midieron directamente al aire y a presión atmosférica. Cada espectro LIBS se obtuvo con seis pulsos láser: los tres primeros pulsos fueron usados



para limpiar la superficie de la muestra y los tres siguientes fueron promediados para la obtención de un único espectro. Para cada muestra se obtuvieron un total de 50 espectros moviendo el posicionador 0.25 mm después de cada 6 pulsos para evitar la adquisición de espectros en posiciones previamente irradiadas. Solo en el caso del fémur derecho de todos los individuos donde se obtuvieron dos conjuntos de 50 espectros cada uno. El primero de ellos (librería de entrenamiento) se usó para la estimación de los modelos y el segundo conjunto para realizar la predicción (librería de test) y con ello la validación interna de los modelos.

Para evitar posibles fluctuaciones espectrales, todos los espectros fueron normalizados con la línea de Ca a 393.37nm. La Fig. 6.2.1 muestra un espectro LIBS típico normalizado, donde se han identificado las principales emisiones atómicas e iónicas.

La capacidad de generalización de los modelos de clasificación se estimó a través de la predicción de muestras óseas no incluidas en el entrenamiento. El estudio de la capacidad de generalización de los modelos no solo tiene cuenta de la capacidad del modelo para clasificar muestras desconocidas sino que también sirve como indicativo del posible sobreajustes de los modelos. Una buena capacidad de generalizar no es el único indicativo de un buen sistema de clasificación, sino que es necesario una validación externa independiente para poder asegurarlo. En nuestro caso, dicha validación se ha realizó a través de la medida de la robustez, eliminado un individuo del entrenamiento cada vez y observando la predicción del modelo a dicho individuo.

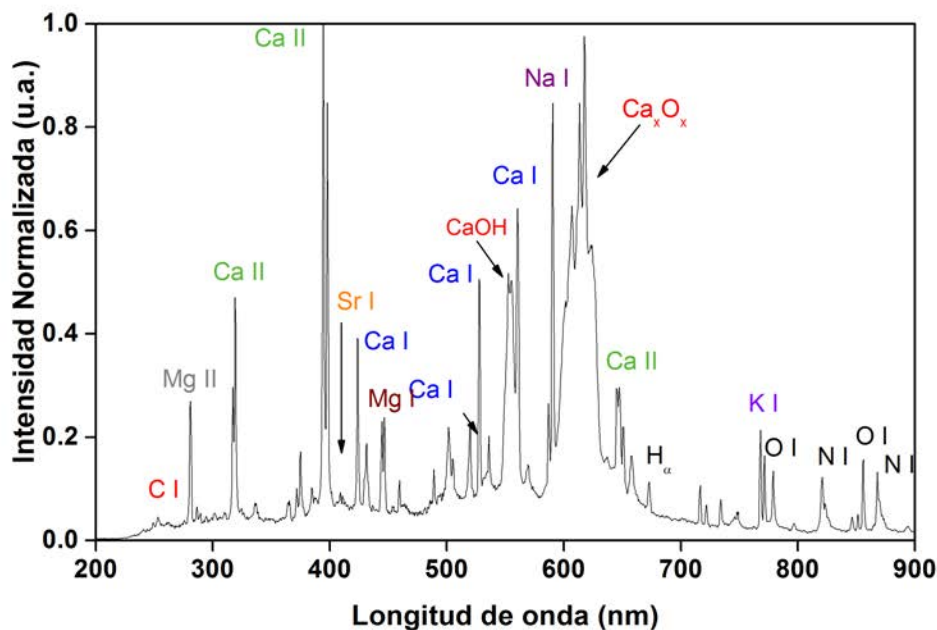


Figura 6.2.1: Espectro típico normalizado de una muestra de hueso con las principales líneas de emisión de átomos neutros, iones y moléculas.

Selección de los intervalos de longitud de onda

En el caso de los espectros LIBS el valor de intensidad a cada longitud de onda es considerado una variable y por lo tanto cada espectro contiene 1800 variables. Para todos los métodos quimiométricos se aplicó una reducción en el número de variables. La reducción de variables debe realizarse de modo que las nuevas variables reproduzcan la parte significativa del espectro sin pérdida de información relevante. Los intervalos de longitud de onda seleccionados 390-410, 420-480 y 516-532 nm tienen en cuenta las principales líneas de emisión de elementos representativos de la muestra como son Ca, Mg y Sr generando un total de 195 variables. Solo en el caso del análisis discriminante lineal (LDA) el número de variables hubo de ser reducido hasta 50 variables, esto es debido a que el número de variables ha de ser como máximo igual al número de objetos con los que se va a estimar el modelo de clasificación. Para reducir de 195 a 50 el



número de variables se aplicó un análisis de componentes principales y se seleccionaron aquellas que tenían una mayor repercusión sobre el modelo (mayor leverage). Por otra parte son muchos los autores que afirman que el uso de todas las variables para llevar a cabo el análisis es innecesario en la mayoría de ocasiones o incluso puede ir en detrimento del rendimiento de los modelos de clasificación.²²⁶ Esto es debido a que para los modelos lineales todas las variables son consideradas igualmente importantes y por tanto pone al mismo nivel variables con información significativa y variables relacionadas con el ruido instrumental. En el caso concreto de las redes neuronales la selección de variables permite ajustar su poder de clasificación. El uso del espectro completo dota a la red de una gran capacidad de discriminación y por tanto es capaz de encontrar diferencias entre muestras del mismo individuo, reduciendo la capacidad de generalizar del modelo. La selección de un número reducido de variables describe la tendencia del individuo, lo que es más importante que la tendencia de un único hueso. En este estudio se ha hecho balance entre el poder de discriminación y la capacidad de generalizar a través de la reducción de las variables.

Análisis de datos

Todos los análisis se realizaron usando Matlab (Mathworks, 2012a y 2014a), a excepción del análisis SVM que se realizó usando *The Unscrambler* (CAMO 10.0.1). PCA, BLR y NN fueron desarrollados en subrutinas desarrolladas en el grupo. SIMCA, LDA, CART y PLS-DA se llevaron a cabo usando un software libre del grupo de quimiometría de Milán desarrollado por Ballabio y colaboradores.²²⁷

6.2.4 Resultados y discusión

Para evaluar el rendimiento en la clasificación de los diferentes métodos supervisados aplicados así como la frecuencia relativa de clasificaciones correctas e incorrectas se utilizó el valor de la exactitud (A):^{69, 228}

$$A = \frac{TP+TN}{TP+TN+FP+FN} \quad (\text{Eq.6.1})$$

En un proceso de clasificación la salida puede ser positiva (P) o negativa (N) y por tanto se pueden dar 4 posibles situaciones:

- Verdadero positivo (TP): caso positivo clasificado como positivo.
- Verdadero negativo (TN): caso negativo clasificado como negativo.
- Falso positivo (FP): caso negativo clasificado como positivo.
- Falso negativo (FN): caso positivo clasificado como negativo.

A partir de estos valores se puede definir una estadística más elaborada que permita el cálculo de parámetros como la sensibilidad, especificidad, precisión y exactitud.

La tasa de éxito de la clasificación, calculada como el porcentaje de huesos correctamente clasificados sobre el total, se usó como medida para estimar la sensibilidad y robustez de los modelos de clasificación. En este estudio la sensibilidad (validación interna) se usó para evaluar la capacidad del modelo en clasificar muestras con las cuales dicho modelo fue estimado, en otras palabras espectros diferentes del mismo hueso de entrenamiento (fémur derecho). La capacidad de generalizar estima la tasa de clasificaciones correctas del modelo cuando es interrogado con muestras que pertenecen a una clase del modelo pero que no han sido usadas en el entrenamiento (huesos diferentes al fémur derecho). Por otra parte, también se estudió la robustez de los modelos definida como la capacidad de detectar como desconocido muestras que no pertenecen a ninguna de las clases utilizadas en la elaboración del modelo.¹¹⁵ Por tanto

la robustez del modelo es una validación externa independiente. En este caso se utilizó una validación de tipo “leave one out” para el estudio de la robustez de los modelos. Para considerar un hueso correctamente clasificado, el modelo debe predecir la pertenencia a cierta clase con un valor superior a un 80 % de correlación espectral (SC) (Eq. 6.2) y menor al 20 % para cualquier otra clase. Para clasificar un hueso como desconocido la correlación espectral ha de ser menor al 80 % para todas las clases.

$$SC = \frac{100}{N} \sum_i \delta_i \quad (\text{Eq. 6.2})$$

donde δ_i es el número de espectros correctamente clasificados y N el número total de espectros predichos.

Resultados PCA

La Fig. 6.2.2 muestra el conjunto de espectros de entrenamiento (cincuenta espectros del fémur de cada individuo) de acuerdo a los valores de las puntuaciones obtenidos en el análisis de componentes principales. En particular, se muestran los valores de la segunda componente principal (CP) frente a la primera. Las dos primeras CPs acumulan el 82.8 % de la varianza total del espacio muestral. Cada clase se muestra con un color diferente para una mejor visualización. Las elipses de confianza de cada clase se calcularon estadísticamente utilizando el código libre SAISIR²²⁹, considerando un $\alpha = 0.05$, mostrando como los individuos 1 y 2 presentan una buena separación del resto de los individuos. Por otro lado, los espectros pertenecientes a los individuos 3, 4 y 5 son difíciles de distinguir debido a la superposición, especialmente en el caso de los individuos 4 y 5, donde los espectros se encuentran completamente mezclados. Como se puede observar en los individuos 1 y 5 contienen algunos factores dominantes en las primeras CPs, mientras que para el resto de los individuos de los primeros CPs no reflejan dichos factores discriminantes.

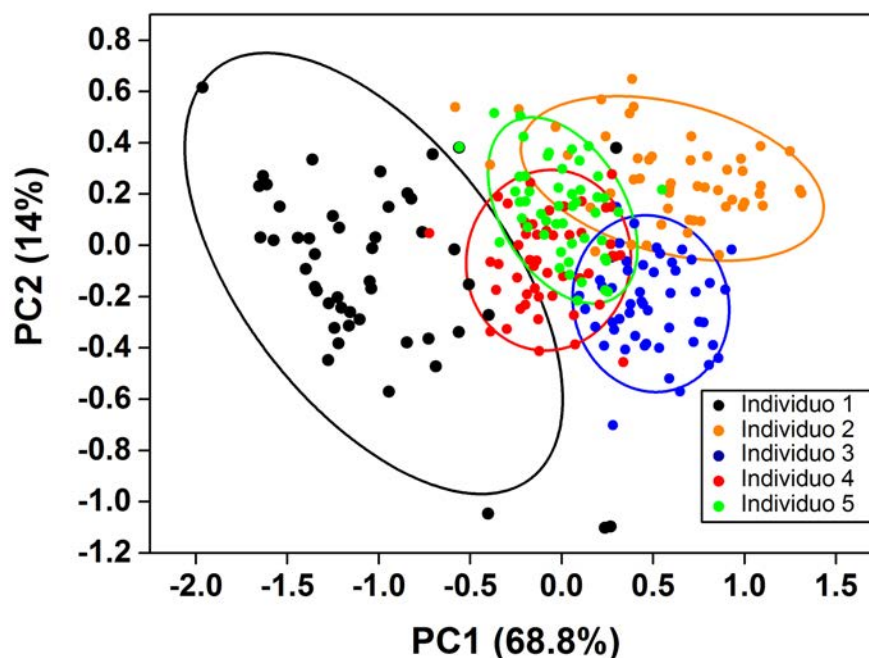


Figura 6.2.2: *Análisis de Componentes Principales para los huesos de entrenamiento (50 espectros x 5 individuos). Se representan los scores de la primera y segunda componente principal, representando un 82.8 % de la varianza total.*

Resultados SIMCA

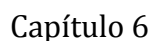
Para la estimación del modelo de clasificación SIMCA, se utilizaron las librerías espectrales de entrenamiento del fémur derecho de cada individuo. Considerando un total de 250 espectros (50 espectros por individuo), para crear el modelo. Para obtener el número adecuado de CPs de cada una de las clases se aplicó un procedimiento de validación de cruce “*leave-one-out*”. El número óptimo de CPs obtenidos para todos los individuos fue de entre 3 y 6. Una vez el modelo ha sido generado, se calculó la distancia entre cada uno de los espectros y el centro de todas las clases. Los espectros son clasificados como pertenecientes a una clase de acuerdo a sus similitudes espectrales con los espectros utilizados en la estimación del modelo. Un espectro desconocido es asignado sólo a la/las clase/s con las que tiene una alta probabilidad de

pertenencia. Un espectro puede no ser asignado como perteneciente a ninguna de las clases cuando su varianza residual excede el límite superior fijado para todas las clases. Dicho límite se fijó siguiendo el método descrito por Sirven et al¹¹⁵ asumiendo que las distancias de muestra de entrenamiento al centro de la correspondiente clase sigue una distribución normal. Un espectro se consideró no clasificado cuando la probabilidad de pertenencia para todas las clases fue de menor del 5 %. La Tabla 6.2.2 muestra los resultados de la predicción de SIMCA para los 25 huesos utilizados en el estudio.

En términos de sensibilidad, 3 de los 5 fémures derechos considerados fueron clasificados correctamente. Desde el punto de vista de la generalización, el modelo sólo fue capaz de clasificar correctamente 2 de los 25 restos. Sólo 5 de 25 huesos fueron clasificados correctamente dando un valor de 20 % en la tasa de éxito de la clasificación. Estos resultados indican una baja sensibilidad y una baja capacidad de generalización del modelo SIMCA en la discriminación y agrupamiento de individuos. En cuanto al test de robustez, todas las muestras fueron correctamente clasificadas como desconocidas, excepto el hueso 31, que fue clasificado erróneamente como pertenecientes al individuo 1. Por lo tanto, la robustez del modelo fue de 96 %. Considerando la tasa de éxito de la clasificación y la robustez se obtuvo un valor de exactitud del 58 % para el modelo de SIMCA.

Resultados PLS-DA

El estudio de la clasificación usando PLS-DA se realizó usando dos algoritmos diferentes, PLS1 y PLS2. Finalmente y debido a la realización de un estudio múltiple (más de dos clases consideradas) el algoritmo PLS2 fue elegido en vista a sus resultados comparados con los obtenidos mediante el algoritmo PLS1. La calibración del modelo



Los resultados de la predicción del modelo PLS-DA se muestran en la Tabla 6.2.3. En este caso todos los huesos utilizados en la calibración del modelo son clasificados correctamente, mostrando por tanto una alta sensibilidad. Sin embargo, la clasificación del resto de huesos no fue satisfactoria y sólo 10 de los 25 huesos fueron clasificados correctamente, mostrando una baja capacidad de generalización y produciendo un 40 % en la tasa de éxito de la clasificación. Por otra parte, los valores del test de robustez (76 %) muestran una robustez moderada, teniendo en cuenta que las muestras pertenecientes al individuo 3 son clasificadas como pertenecientes al individuo 1 y no como desconocido, generando así una alta tasa de falsos positivos. Teniendo en cuenta ambos test, la exactitud de este modelo fue del 58 %.

**Resultados LDA**

Como se ha comentado anteriormente, en el caso del modelo de LDA el número de variables se redujo a 50 (número de muestras) mediante un modelo PCA. El modelo se estimó utilizando los conjuntos de datos de entrenamiento de todos los individuos de forma simultánea. El modelo LDA genera $N-1$ función discriminante, siendo N es el número de clases. Una vez el modelo ha sido generado, las librerías de test de todos los individuos fueron introducidas en el modelo para obtener los resultados de predicción (Tabla 6.2.4). LDA clasificó correctamente 16 de los 25 huesos obteniendo una tasa de éxito en la clasificación del 64 % y una robustez del 20 %. Es importante mencionar que este modelo no puede predecir muestras no pertenecientes a ninguna clase, se trata de un método de clasificación puro y por tanto todas las muestras siempre son asignadas al menos a una clase. Esto implica que el método no es adecuado para ser aplicado en casos donde haya muestras desconocidas que no pertenecen a ninguna clase. Por otra parte, aunque la tasa de éxito del modelo fue la mayor entre los modelos lineales, produjo una robustez muy baja y con ello una exactitud del 42 %.

Resultados CART

La Tabla 6.2.5 muestra los resultados de predicción del método CART. Para la estimación del modelo, las cinco clases se etiquetaron con variables categóricas numeradas del 1 al 5 correspondientes a cada individuo considerado. El modelo divide los datos de entrenamiento en regiones independientes basándose en decisiones binarias y conservando aquellas que producen un mejor resultado de clasificación. Una desventaja del modelo CART es que puede ser fácilmente sobrestimado, adaptándose en exceso a los datos de entrenamiento. Reduciendo los niveles de crecimiento, se evita el

sobreentrenamiento y se aumenta la capacidad de generalizar. Como en el caso de LDA, CART es una técnica de clasificación pura, lo que implica una división del hiperespacio en diferentes regiones (número de clases) y por lo tanto, todos los objetos se clasifican siempre como pertenecientes a una clase, sin posibilidad de realizar una predicción como no clasificado. Los resultados obtenidos por el método CART muestran una clasificación correcta de 12 de los 25 huesos obtención de 48 % de tasa de éxito de la clasificación y una robustez de 48 % dando lugar a una exactitud global del 48 %.

Resultados BLR

El modelo BLR calcula la probabilidad de clasificación de una muestra como 'negativo' ó 'positiva' de forma binaria. Y_i indican la predicción de la i -ésima de una muestra, cuando $Y_i > 0.8$ la muestra se clasifica como "positiva" y si $Y_i < 0.8$ como negativo. El modelo BLR utiliza una función logística para relacionar las variables con la probabilidad y obtener el vector de los coeficientes de regresión utilizando el máximo de probabilidad. Para el test de sensibilidad, se estimaron cinco modelos diferentes, mientras que para la prueba de robustez se usaron un total de veinte modelos, cubriendo todos los posibles modelos de entrenamiento. Los resultados del método BLR se muestran en la Tabla 6.2.6. La tasa de éxito de la clasificación fue del 68 %, 17 huesos de 25 se clasificaron correctamente con nivel de significación del 5 %. Los resultados del test de robustez muestran que 16 de 25 huesos fueron clasificados correctamente como desconocido, dando un 64 % de robustez. La exactitud global del método fue del 66 %.

Resultados SVM

La Tabla 6.2.7 muestra los resultados de predicción del método de SVM. Para la optimización de los parámetros C y γ del modelo se realizó una búsqueda fina,



evaluando como la modificación de dichos parámetros afectaban al modelo de clasificación. Finalmente se seleccionaron aquellos que producían un mejor ajuste a los datos de entrenamiento y proporcionaban el mínimo error en la validación de cruce. Los valores encontrados fueron 1000 y 0.000606 para los parámetros C y γ , respectivamente, conduciendo a un valor de 97.2 % en la validación de cruce. En el modelo SVM, los datos de entrada son transformados a través de la función Kernel a un espacio de mayores dimensiones, esta transformación normalmente provoca un menor rendimiento del modelo de clasificación, a pesar de ello, el método SVM suele generalizar correctamente. En nuestro caso, para mantener el rendimiento del modelo, se utilizó un número alto de muestras de entrenamiento. Como regla general a mayor dimensionalidad, se requiere mayor número de muestras de entrenamiento. Una vez se llevó a cabo la etapa de entrenamiento, las librerías espectrales de test de todos los huesos fueron introducidos en el modelo para obtener los resultados de la predicción. SVM logra la clasificación correcta de 15 de los 25 huesos obteniendo una tasa de éxito en la clasificación del 60 % y una robustez del 52 %. La exactitud global del método fue del 56 %. Este método permite trabajar con relaciones no lineales entre los datos de entrada, presentando una capacidad de generalización aceptable a pesar del uso de un número de muestras de entrenamiento limitadas.

En cuanto a sus limitaciones, es importante destacar la tediosa labor de encontrar los parámetros óptimos del modelo, especialmente la penalización de error (parámetro C). El método SVM fue inicialmente creado para realizar clasificaciones binarias, el entrenamiento de SVM multiclase, es complejo y difícil de optimizar y por tanto a pesar de sus buenos resultados a día de hoy está lejos de ofrecer sus mejores resultados y se necesita una mayor labor en el diseño de un modelo multiclase óptimo basado en SVM.

Resultados NN

Para la estimación del modelo de NN se utilizó una arquitectura de NN tipo *perceptron* multicapa con conexiones hacia adelante, seleccionando un algoritmo de entrenamiento tipo *back-propagation* (BP). La arquitectura de la NN utilizada ya ha sido descrita en la aplicación anterior y en la siguiente publicación del grupo.⁵⁵ Como en el resto de modelos, el entrenamiento de la red se llevó a cabo con las librerías espectrales de entrenamiento del fémur derecho de todos los individuos. Cada individuo fue asignado con número de identificación en la capa de salida (5 neuronas). Una vez obtenido el entrenamiento del modelo, se introdujeron en el modelo todas las librerías espectrales de test del resto de huesos para obtener su predicción. La Tabla 6.2.8 muestra los resultados de la clasificación de todos los huesos de los cinco individuos considerados. En primer lugar es importante destacar que todas las muestras fueron clasificadas correctamente con una correlación espectral superior a 98 %. Es importante tener en cuenta que igual que para el resto de métodos sólo un hueso de cada individuo se utilizó para la estimación del modelo de NN. Además no solo se ha obtenido un excelente resultado en la sensibilidad del modelo, si no que la red neuronal ha demostrado una alta capacidad de generalizar asignando correctamente todos los huesos con su correspondiente individuo, 100 % de tasa de éxito en la clasificación. En multitud de ocasiones, la capacidad de generalizar de las redes neuronales se pone en entredicho debido a su alto riesgo de ser sobreentrenada. Sin embargo, cuando se estudian y optimizan los parámetros que controlan el entrenamiento, funciones de transferencia, gradientes, algoritmo de entrenamiento, etc, es posible llegar a modelos con una alta sensibilidad sin perder capacidad de generalización. En este estudio, la búsqueda de las condiciones más favorables a resultado en un modelo altamente

confiable, con una excelente capacidad de generalizar y una alta robustez (100 %). Sin falsos positivos ni falsos negativos y proporcionando una precisión global de 100 %.

Comparativa de Métodos

La comparación de los resultados globales de los siete métodos supervisados se muestra en la [Fig. 6.2.3](#). El método SIMCA no permitió la clasificación de espectros no utilizados en la etapa de entrenamiento, mostrando una baja sensibilidad y capacidad de generalización. Sin embargo, presenta una alta capacidad para detectar muestras desconocidas correctamente y produce una alta robustez. En general el método no ofrece la precisión adecuada ni las características necesarias de un modelo de clasificación para agrupar individuos en base a datos espectrales. Estos resultados demuestran que la composición elemental de los huesos, incluso dentro de un mismo individuo, difieren significativamente no permitiendo su discriminación. Un valor similar en términos de exactitud fue observado cuando se aplicó el método PLS-DA, sin embargo, la tasa de éxito en este caso fue el doble que para SIMCA, mostrando por tanto una mayor capacidad de generalización. Por otro lado, el test de robustez para el método PLS-DA presenta una mayor tasas de falsos positivos, en particular para el individuo 3 donde todos los huesos fueron asignados como pertenecientes al individuo 4 y por tanto reduciendo la robustez del modelo hasta el 76 %. Los resultados de LDA son buenos en cuanto a sensibilidad y capacidad de generalización produciendo una tasa de éxito en la clasificación del 64 %. Sin embargo, su mayor desventaja es la imposibilidad de clasificar muestras como desconocidas, en este caso la robustez del método fue solamente del 20 %. Existen algunos trabajos donde se intenta fijar límites en las distancias para la clasificación LDA, sin embargo la aplicación de estos límites ha de ser optimizada para cada caso particular y no existe ningún protocolo establecido. Por



tanto, para casos donde conocemos que todas las muestras pertenecen al menos a una de las clases de entrenamiento el método LDA es una buena opción, siendo un método sencillo y fácilmente implementable. En nuestro caso, los resultados obtenidos con LDA fueron similares a los obtenidos con el método SVM, a pesar de que SVM usa relaciones no lineales.

Por su parte, el método CART mostró un rendimiento modesto en cuanto a la tasa de éxito y robustez, produciendo una exactitud global del 48 %. El algoritmo de los árboles de decisiones es complejo y las decisiones se toman de forma binaria, determinando en cada momento dos alternativas posibles, hasta la construcción final del árbol, con el objetivo de generar grupos homogéneos respecto a la variable que se desea discriminar. Sin embargo, en la bibliografía se pone de manifiesto la inestabilidad de este método y como pequeñas diferencias en los datos de entrenamiento pueden producir modelos de clasificación distintos.

El método BLR mostró los segundos mejores resultados de entre los siete métodos supervisados estudiado, viéndose sólo superado por las redes neuronales. Sin embargo se trata de un método binario y no multiclase como el resto.

En cuanto a las redes neuronales, fueron las que mejores resultados de clasificación produjeron en todos los aspectos estudiados, sensibilidad, capacidad de generalización y robustez, produciendo en todos los casos valores del 100 % de correlación espectral. Por lo tanto, a pesar de las diferencias en la composición elemental de los diferentes huesos de una misma persona, todas las muestras fueron asignadas correctamente. Podemos decir, en vista a los resultados, que la NN es capaz de maximizar las diferencias entre individuos y minimizar la variabilidad dentro de un mismo individuo, permitiendo así la discriminación en base a las librerías espectrales obtenidas por LIBS.

Es muy importante destacar que, aunque nosotros conocemos el origen y la pertinencia de todos los huesos, cuando son introducidos en el modelo son considerados completamente desconocidos. Los resultados de clasificación dependen exclusivamente de la capacidad de generalización de cada método y su robustez. Con este estudio se pretende resolver una problemática actual, en una situación real, donde la identificación de los restos implica un alto coste y no es posible la identificación a través de técnicas convencionales descritas en el ámbito de la antropología forense. Para la aplicación de este procedimiento de trabajo, al menos un hueso de cada individuo debe ser identificado para ser utilizado como referencia. Cualquier hueso que no pertenezca a ninguna de las clases de entrenamiento será clasificado como desconocido y podría ser incluido en el modelo como una nueva clase, aumentando la librería espectral de referencia.

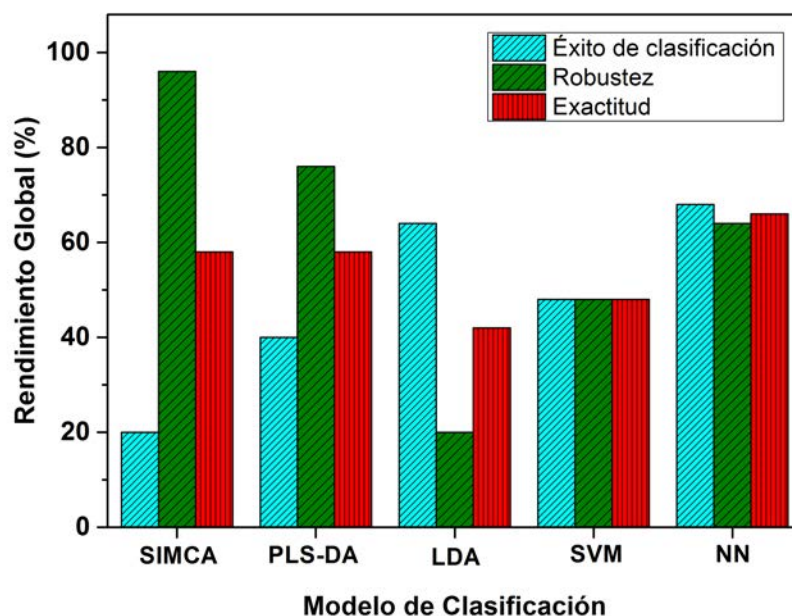


Figura 6.2.3: Rendimiento global de los siete métodos quimiométricos estudiados. Se representa en azul la tasa de éxito de la clasificación, en verde la robustez y en rojo la exactitud.



6.2.5 Conclusiones a la evaluación de métodos quimiométricos supervisados en la discriminación de individuos mediante LIBS

La espectroscopia de ablación láser LIBS y siete métodos quimiométricos supervisados (SIMCA, PLS-DA, LDA, CART, BLR, SVM y NN) han sido estudiados en una aplicación real para la discriminación rápida y robusta de restos humanos, basado en las huellas espectrales obtenidas a través de restos óseos. Los espectros LIBS se encuentran comúnmente afectados por efectos no lineales como la autoabsorción, fluctuaciones en la energía del pulso láser y efecto del ruido de fondo, entre otros. Por ello, los métodos no lineales, como las redes neuronales, resultan ser más adecuados para llevar a cabo la clasificación.

Por último, se ha demostrado el potencial de los modelos de NN para modelar relaciones no lineales complejas entre las variables. Aunque los modelos de NN puede ser tediosos a la hora de su implementación, los resultados obtenidos demuestran que la combinación entre LIBS y NN ofrece ventajas significativas a la hora de llevar a cabo una estudio cualitativo.



Samuel Moncayo Martín

Tabla 6.2.2: Resultados clasificación SIMCA

Predicción grupo pertenencia (Correlación espectral, %)													
Test Identificación								Test Robustez					
Hueso	Ind. 1	Ind. 2	Ind. 3	Ind. 4	Ind. 5	Sin clasificar	Correctamente clasificado	Ind. 1	Ind. 2	Ind. 3	Ind. 4	Ind. 5	Correctamente como desconocido
11	100	0	0	0	0	0	✓	*	0	0	0	0	✓
12	48	0	0	0	16	44	×	*	0	0	0	0	✓
13	0	0	0	0	0	100	×	*	0	0	0	0	✓
14	96	0	0	0	0	4	✓	*	0	0	0	0	✓
15	0	0	0	0	0	100	×	*	0	0	0	0	✓
21	4	100	0	0	0	0	✓	0	*	0	0	0	✓
22	8	72	0	0	0	20	×	8	*	0	0	0	✓
23	8	72	0	0	0	20	×	8	*	0	0	0	✓
24	4	100	0	0	0	0	✓	8	*	0	0	0	✓
25	92	24	0	40	0	4	×	16	*	0	0	0	✓
31	8	4	100	40	0	0	×	96	0	*	0	0	×
32	8	32	100	0	0	0	×	24	0	*	0	0	✓
33	16	28	80	48	0	4	×	16	28	*	0	0	✓
34	24	16	100	44	0	0	×	40	0	*	4	0	✓
35	12	24	56	44	0	24	×	0	24	*	4	0	✓
41	40	12	56	100	0	0	×	92	28	0	*	0	✓
42	44	28	92	32	0	0	×	56	28	80	*	0	✓
43	100	0	4	56	0	0	×	32	16	28	*	0	✓
44	60	20	80	68	0	0	×	76	20	56	*	0	✓
45	80	12	72	100	0	0	×	68	12	8	*	0	✓
51	0	0	0	0	92	8	✓	16	0	0	0	*	✓
52	0	0	0	0	28	72	×	0	0	0	0	*	✓
53	0	0	0	0	60	40	×	8	0	0	0	*	✓
54	0	40	0	0	48	48	×	20	0	0	0	*	✓
55	4	8	0	0	40	64	×	32	0	0	0	*	✓

* clase eliminada del entrenamiento.



Samuel Moncayo Martín

Tabla 6.2.3: Resultados clasificación PLS-DA

Predicción grupo pertenencia (Correlación espectral, %)													
Test Identificación								Test Robustez					
Hueso	Ind. 1	Ind. 2	Ind. 3	Ind. 4	Ind. 5	Sin clasificar	Correctamente clasificado	Ind. 1	Ind. 2	Ind. 3	Ind. 4	Ind. 5	Correctamente como desconocido
11	88	0	4	4	0	0	✓	*	0	8	0	40	✓
12	92	0	68	0	52	0	×	*	0	0	64	0	✓
13	0	0	88	20	0	0	×	*	64	0	100	0	✓
14	64	0	4	0	0	0	×	*	12	0	0	0	✓
15	0	0	20	4	0	76	×	*	0	0	0	0	✓
21	0	88	0	4	0	0	✓	0	*	0	0	0	✓
22	24	0	16	0	0	60	×	0	*	24	0	20	✓
23	24	0	20	0	0	56	×	0	*	28	0	20	✓
24	16	4	8	0	0	72	×	0	*	12	0	16	✓
25	0	0	0	4	0	96	×	0	*	0	100	0	×
31	0	0	96	0	0	4	✓	0	0	*	92	0	×
32	0	0	96	0	0	4	✓	0	0	*	96	0	×
33	0	0	4	96	0	0	×	0	0	*	96	0	×
34	0	0	88	4	0	20	✓	0	0	*	100	0	×
35	0	0	0	96	0	4	×	0	0	*	96	0	×
41	0	4	0	82	0	0	✓	20	0	0	*	0	✓
42	4	0	0	8	0	64	×	8	16	40	*	0	✓
43	0	0	0	92	0	8	✓	0	4	40	*	0	✓
44	0	0	0	96	0	4	✓	28	20	20	*	0	✓
45	0	0	0	96	0	4	✓	4	0	32	*	0	✓
51	0	0	0	0	100	0	✓	0	0	0	0	*	✓
52	8	0	0	0	52	40	×	0	0	0	0	*	✓
53	0	0	0	0	56	44	×	0	0	0	0	*	✓
54	56	0	0	0	4	40	×	0	0	0	0	*	✓
55	80	0	0	0	4	16	×	0	0	0	0	*	✓

* clase eliminada del entrenamiento.



Samuel Moncayo Martín

Tabla 6.2.4: Resultados clasificación LDA

Predicción grupo pertenencia (Correlación espectral, %)													
Test Identificación								Test Robustez					
Hueso	Ind. 1	Ind. 2	Ind. 3	Ind. 4	Ind. 5	Sin clasificar	Correctamente clasificado	Ind. 1	Ind. 2	Ind. 3	Ind. 4	Ind. 5	Correctamente como desconocido
11	100	0	0	0	0	E	✓	*	0	4	96	0	×
12	100	0	0	0	0	E	✓	*	36	8	52	4	✓
13	20	0	0	0	80	E	×	*	52	4	44	0	✓
14	100	0	0	0	0	E	✓	*	60	0	40	0	✓
15	4	0	0	0	96	E	×	*	0	0	0	100	×
21	0	100	0	4	0	E	✓	4	*	96	0	0	×
22	44	0	52	0	4	E	×	0	*	100	0	0	×
23	44	0	56	0	0	E	×	0	*	0	100	0	×
24	36	4	60	0	0	E	×	0	*	100	0	0	×
25	0	40	4	66	0	E	×	0	*	0	100	0	×
31	0	0	100	0	0	E	✓	44	66	*	0	0	✓
32	0	0	100	0	0	E	✓	0	16	*	84	0	×
33	0	0	0	100	0	E	×	0	0	*	100	0	×
34	0	0	100	0	0	E	✓	4	20	*	76	0	✓
35	0	0	4	96	0	E	×	0	0	*	100	0	×
41	0	0	0	100	0	E	✓	0	0	100	*	0	×
42	0	0	72	28	0	E	×	0	4	96	*	0	×
43	0	0	0	100	0	E	✓	0	8	92	*	0	×
44	0	0	0	100	0	E	✓	0	8	92	*	0	×
45	0	0	0	100	0	E	✓	0	0	100	*	0	×
51	0	0	0	0	100	E	✓	100	0	0	0	*	×
52	0	0	0	0	100	E	✓	100	0	0	0	*	×
53	0	0	0	0	100	E	✓	100	0	0	0	*	×
54	0	0	0	0	100	E	✓	100	0	0	0	*	×
55	0	0	0	0	100	E	✓	100	0	0	0	*	×

E Resultado no válido para LDA

* clase eliminada del entrenamiento.



Samuel Moncayo Martín

Tabla 6.2.5: Resultados clasificación CART

Predicción grupo pertenencia (Correlación espectral, %)													
Test Identificación								Test Robustez					
Hueso	Ind. 1	Ind. 2	Ind. 3	Ind. 4	Ind. 5	Sin clasificar	Correctamente clasificado	Ind. 1	Ind. 2	Ind. 3	Ind. 4	Ind. 5	Correctamente como desconocido
11	98	2	0	0	0	E	✓	*	0	28	32	50	✓
12	98	0	0	2	0	E	✓	*	0	8	92	0	×
13	0	100	0	0	0	E	×	*	100	0	0	0	×
14	94	0	2	4	0	E	✓	*	20	28	52	0	✓
15	14	86	0	0	0	E	×	*	72	28	0	0	✓
21	0	94	0	6	0	E	✓	0	*	0	88	12	×
22	4	76	6	0	4	E	×	8	*	84	0	8	×
23	4	76	6	0	4	E	×	8	*	84	0	8	×
24	0	96	0	4	0	E	✓	0	*	92	8	0	×
25	2	0	0	98	0	E	×	8	*	0	82	0	×
31	8	0	92	0	0	E	✓	0	0	*	40	60	✓
32	0	0	100	0	0	E	✓	0	0	*	20	80	×
33	0	0	68	32	0	E	×	0	0	*	24	76	✓
34	0	0	100	0	0	E	✓	0	4	*	40	66	✓
35	0	0	68	32	0	E	×	0	0	*	16	86	×
41	44	0	0	66	0	E	×	100	0	0	*	0	×
42	0	0	82	8	0	E	×	12	0	88	*	0	×
43	0	0	0	100	0	E	✓	60	24	16	*	0	✓
44	0	0	28	72	0	E	×	36	0	64	*	0	✓
45	4	0	0	96	0	E	✓	84	4	12	*	0	×
51	4	4	0	0	92	E	✓	0	32	68	0	*	✓
52	12	0	0	0	88	E	✓	0	28	72	0	*	✓
53	80	0	0	0	20	E	×	0	12	88	0	*	×
54	44	4	12	0	40	E	×	0	52	48	0	*	✓
55	36	0	20	0	44	E	×	0	76	24	0	*	✓

E Resultado no válido para CART

* clase eliminada del entrenamiento.



Samuel Moncayo Martín

Tabla 6.2.6: Resultados clasificación Regresión Logística

Predicción grupo pertenencia (Correlación espectral, %)													
Test Identificación								Test Robustez					
Hueso	Ind. 1	Ind. 2	Ind. 3	Ind. 4	Ind. 5	Sin clasificar	Correctamente clasificado	Ind. 1	Ind. 2	Ind. 3	Ind. 4	Ind. 5	Correctamente como desconocido
11	96	0	8	0	4	E	✓	*	0	40	40	76	✓
12	100	0	4	0	0	E	✓	*	8	4	96	4	x
13	8	0	100	0	40	E	x	*	80	92	0	4	x
14	60	0	44	4	0	E	x	*	8	0	96	0	x
15	12	4	100	0	36	E	x	*	100	100	0	0	✓
21	0	96	4	32	0	E	✓	24	*	4	92	4	x
22	0	56	88	0	0	E	x	32	*	28	24	4	✓
23	0	56	88	0	0	E	x	32	*	28	24	4	✓
24	0	40	64	0	0	E	x	20	*	20	32	0	✓
25	40	80	0	96	0	E	x	60	*	0	92	0	✓
31	0	0	96	0	0	E	✓	12	0	*	100	8	x
32	0	0	100	0	0	E	✓	0	0	*	100	4	x
33	0	60	0	68	0	E	x	0	80	*	56	0	✓
34	0	0	100	0	0	E	✓	4	0	*	100	0	x
35	0	44	0	84	0	E	x	0	92	*	60	0	✓
41	0	0	0	100	0	E	✓	16	0	100	*	0	x
42	0	4	0	72	0	E	x	0	32	68	*	0	✓
43	0	0	0	100	0	E	✓	0	24	100	*	0	✓
44	0	0	0	100	0	E	✓	0	44	64	*	0	✓
45	0	0	0	100	0	E	✓	0	4	96	*	0	x
51	0	0	0	12	100	E	✓	68	8	40	0	*	✓
52	0	8	0	0	100	E	✓	72	96	28	0	*	✓
53	0	0	0	0	100	E	✓	8	96	84	0	*	✓
54	0	0	0	12	100	E	✓	0	72	80	0	*	✓
55	0	0	0	0	100	E	✓	0	88	60	0	*	✓

E Resultado no válido para CART

* clase eliminada del entrenamiento.



Samuel Moncayo Martín

Tabla 6.2.7: Resultados clasificación SVM

Predicción grupo pertenencia (Correlación espectral, %)													
Test Identificación								Test Robustez					
Hueso	Ind. 1	Ind. 2	Ind. 3	Ind. 4	Ind. 5	Sin clasificar	Correctamente clasificado	Ind. 1	Ind. 2	Ind. 3	Ind. 4	Ind. 5	Correctamente como desconocido
11	96	0	4	0	0	0	✓	*	36	12	28	24	✓
12	100	0	0	0	0	0	✓	*	60	0	40	0	✓
13	24	64	0	0	12	0	×	*	100	0	0	0	×
14	100	0	0	0	0	0	✓	*	64	0	36	0	✓
15	100	0	0	0	0	0	✓	*	100	0	0	0	×
21	0	96	0	4	0	0	✓	40	*	0	8	52	✓
22	8	88	0	4	0	0	✓	32	*	48	20	0	✓
23	0	96	0	4	0	0	✓	32	*	48	0	0	✓
24	0	84	0	16	0	0	✓	12	*	12	76	0	✓
25	4	0	0	96	0	0	×	4	*	0	96	0	×
31	0	0	100	0	0	0	✓	0	4	*	92	4	×
32	0	0	100	0	0	0	✓	0	16	*	0	84	×
33	0	0	68	32	0	0	×	0	28	*	0	72	✓
34	0	0	100	0	0	0	✓	0	32	*	68	0	✓
35	4	0	64	32	0	0	×	4	12	*	84	0	×
41	0	0	0	100	0	0	✓	100	0	0	*	0	×
42	0	0	80	20	0	0	×	96	4	4	*	0	×
43	0	0	0	100	0	0	✓	100	0	0	*	0	×
44	4	0	40	56	0	0	×	92	8	0	*	0	×
45	68	0	0	32	0	0	×	100	0	0	*	0	×
51	0	0	0	0	100	0	✓	0	92	8	0	*	×
52	0	4	0	0	96	0	✓	0	8	76	16	*	✓
53	28	0	8	0	64	0	×	0	16	8	76	*	✓
54	8	4	36	0	52	0	×	0	36	40	24	*	✓
55	24	8	20	0	48	0	×	0	16	48	36	*	✓

* clase eliminada del entrenamiento.



Samuel Moncayo Martín

Tabla 6.2.8: Resultados clasificación NN

Predicción grupo pertenencia (Correlación espectral, %)													
Test Identificación								Test Robustez					
Hueso	Ind. 1	Ind. 2	Ind. 3	Ind. 4	Ind. 5	Sin clasificar	Correctamente clasificado	Ind. 1	Ind. 2	Ind. 3	Ind. 4	Ind. 5	Correctamente como desconocido
11	100	0	0	0	0	0	✓	*	0	0	0	0	✓
12	100	0	0	0	0	0	✓	*	0	0	0	0	✓
13	100	0	0	0	0	0	✓	*	0	0	0	0	✓
14	100	0	0	0	0	0	✓	*	0	0	0	0	✓
15	100	0	0	0	0	0	✓	*	0	0	0	0	✓
21	0	100	0	0	0	0	✓	0	*	0	0	0	✓
22	0	98	0	0	0	2	✓	0	*	0	0	0	✓
23	0	100	0	0	0	0	✓	0	*	2	0	0	✓
24	0	100	0	0	0	0	✓	8	*	0	20	8	✓
25	0	98	0	0	0	2	✓	12	*	0	0	12	✓
31	0	0	100	0	0	0	✓	0	0	*	0	0	✓
32	0	0	98	0	0	2	✓	0	0	*	0	0	✓
33	0	0	98	0	0	2	✓	0	0	*	0	0	✓
34	0	0	100	0	0	0	✓	0	0	*	0	0	✓
35	0	0	100	0	0	0	✓	0	0	*	0	0	✓
41	0	0	0	100	0	0	✓	0	0	0	*	0	✓
42	0	0	0	98	0	2	✓	0	0	0	*	0	✓
43	0	0	0	100	0	0	✓	0	0	0	*	0	✓
44	0	0	0	98	0	2	✓	0	0	0	*	0	✓
45	0	0	0	100	0	0	✓	0	0	0	*	0	✓
51	0	0	0	0	100	0	✓	0	4	0	0	*	✓
52	0	0	0	0	98	2	✓	0	24	0	0	*	✓
53	0	0	0	0	98	2	✓	0	48	0	0	*	✓
54	0	0	0	0	100	0	✓	0	0	0	0	*	✓
55	0	0	0	0	98	2	✓	0	4	0	0	*	✓

* clase eliminada del entrenamiento.

Agradecimientos:

Agradezco sinceramente la ayuda prestada para la elaboración de este trabajo al resto de autores de la publicación S. Manzoor, T. Ugidos, F. Navarro-Villoslada y J. O. Cáceres y a la Universidad Complutense de Madrid por el proyecto (CCG10-UCM/PPQ-4713).



Capítulo 7:

LIBS aplicado al análisis medioambiental

La conciencia ambiental puede definirse como la percepción que tiene el ser humano sobre su impacto sobre el entorno. En otras palabras, entender cómo influyen las acciones de cada día en el medio ambiente y como esto afecta el futuro de nuestro entorno.

Afortunadamente, la conciencia ambiental, a día de hoy, está en continuo crecimiento y los problemas ambientales son cada vez más estudiados con el objetivo de paliar sus efectos o en el mejor de los casos, incluso eliminarlos. Los problemas ambientales se refieren a situaciones ocasionadas por actividades, procesos o comportamientos humanos, económicos, sociales, culturales y políticos, que trastornan el entorno y ocasionan impactos negativos directos sobre el medioambiente, e indirectamente sobre la economía y la sociedad.

A lo largo del siglo pasado, el calentamiento global, ocasionado por las actividades humanas que emiten dióxido de carbono, ha generado un aumento aproximado de 0.8°C sobre la temperatura media del planeta. En los océanos, este cambio ha sido de aproximadamente 0.1°C.^{230, 231} Este calentamiento se ha producido desde la superficie hasta una profundidad de alrededor de 700 metros, zona donde habita la mayoría de la fauna y flora marina.

Sin duda los arrecifes coralinos ralentizan su crecimiento, haciéndose más propensos a contraer enfermedades y pudiendo causar incluso la extinción masiva de dichos arrecifes.



Por tanto, es necesario el uso de indicadores que permitan estudiar los cambios climáticos para evaluar su efecto sobre el medioambiente.

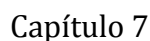
El clima terrestre sufre variaciones naturales, nuestro clima actual difiere mucho del que dominaba hace 100 millones de años, cuando los dinosaurios habitaban el planeta o incluso de lo que era hace 18000 años, cuando la mayor parte del hemisferio norte se encontraba completamente cubierto de hielo. Por tanto, es fácil de imaginar que el clima en el futuro será diferente. Sin embargo, además de las causas naturales, como las fluctuaciones en la órbita terrestre, hay que añadir una nueva fuente de variación importante como es la actividad humana.

Para poder predecir el clima del futuro, se usan modelos matemáticos que permiten realizar una simulación de los cambios climáticos. Sin embargo, dicha modelización, para ser más realista, necesita tener en cuenta los cambios ocurridos en el clima pasado y es ahí donde cobra importancia la paleoclimatología. Con toda esta información se pueden evaluar los cambios y estudiar nuevos métodos para amortiguar de la mejor manera posible los efectos desfavorables.

Los paleoclimatólogos basan su trabajo en el estudio de los llamados “*proxy*” o indicadores naturales, para inferir como fueron las condiciones climáticas en el pasado.

En nuestro estudio se han utilizado como proxy dos tipos de muestras diferentes, moluscos y corales.

Los corales, principalmente constituidos a partir de carbonato de calcio, contienen importante información relacionada con relaciones isotópicas y trazas metálicas que pueden ser usadas para determinar la temperatura del agua, en la que el coral vivió. Estos registros de temperatura son usados para reconstruir el clima durante el periodo de vida del coral, ya sean corales actuales o corales fósiles. Los corales fósiles



185

Resumen

En este trabajo se ha evaluado una especie de molusco *Phorcus lineatus* como proxy medioambiental. Se ha medido por primera vez mediante Laser Induced Breakdown Spectroscopy (LIBS) la relación Mg/Ca en su concha calcificada con el objetivo de inferir información paleoclimática y patrones de estacionalidad de los grupos humanos prehistóricos.

Todas las muestras se obtuvieron con los permisos pertinentes en las costas del mar Cantábrico. Los resultados obtenidos fueron comparados con las temperaturas instrumentales de la superficie del mar (SST) y con medidas de isótopos de oxígeno ($\delta^{18}\text{O}$). Las medidas se realizaron en dos biominerales diferentes presentes en la concha (aragonito y calcita) dando como resultado una correlación entre las relaciones Mg/Ca y la SST de $R^2 = 0.43$ y 0.44 , respectivamente. Se estudiaron también secuencias largas de tres conchas recogidas en otoño de 2012. Los resultados muestran variaciones en la proporción de Mg/Ca relacionado con cambios estacionales en la SST durante todo el año y una buena correlación con las medidas de $\delta^{18}\text{O}$ en dos conchas ($R^2 = 0.70$ y 0.65 , respectivamente).

7.1.1 Introducción al análisis de muestras de moluscos

Las tareas de reconstrucción del clima y la variabilidad medioambiental en épocas pasadas dependen fundamentalmente de la capacidad de obtener proxies que permitan una alta resolución climática. Muchos de ellos están basados en parámetros geoquímicos y por ello, para obtener información útil, es necesario aplicar técnicas analíticas eficientes que permitan obtener un registro geológico, paleontológico y arqueológico de calidad.²³³ Los restos biológicos encontrados en sitios arqueológicos no solo nos permiten obtener información de climas pasados, sino que además son piezas



clave para entender factores de la actividad humana de homínidos prehistóricos. Existen multitud de restos biológicos con una gran importancia desde el punto de vista arqueológico, entre todos ellos, se han elegido los restos de moluscos ya se encuentran en grandes acumulaciones (concheros) como resultado de la actividad humana pasada.²³⁴ Las conchas de estos moluscos son excelentes archivos de las condiciones ambientales en las cuales vivieron, quedando reflejados cambios debidos a la estacionalidad y que sirven como indicadores del comportamiento de los seres humanos, revelando factores como la temporada de captura o estrategias de supervivencia.²³⁵ Los estudios paleoclimáticos están basados fundamentalmente en el análisis de dichas conchas a través de estudios isotópicos, fundamentalmente se usa el valor de $\delta^{18}\text{O}$, debido a su buena correlación con la temperatura de la superficie del mar.²³⁶⁻²³⁹ En menor medida se han utilizado otros datos geoquímicos como las relaciones elementales de metales traza en los carbonatos biogénicos que forman la concha de los moluscos. La temperatura de la superficie del agua del mar está en equilibrio con la incorporación de elementos químicos en los carbonatos. Sin embargo dicha incorporación no sólo depende de factores ambientales, sino que también está altamente influenciada por factores no ambientales, generalmente de tipo biológico. Las capas de precipitación son en ocasiones de tamaño micrométrico, siendo difícil la realización de los estudios a lo largo del eje de crecimiento de los moluscos.

Siguiendo en esta dirección, en este trabajo hemos propuesto una nueva metodología para el análisis de las relaciones elementales Mg/Ca en muestras de moluscos. En este caso hemos seleccionado como objeto de estudio la especie *Phorcus Lineatus*, un gasterópodo abundante en las costas actuales y con una importante contribución en concheros arqueológicos. El estudio de la relación Mg/Ca en otros moluscos ya se ha

realizado por varios autores, tanto en especies en libertad como en cautividad y se han encontrado resultados significativamente diferentes según la especie estudiada.

Se ha encontrado una alta correlación entre la relación Mg/Ca con la temperatura cuando se ha estudiado moluscos como *Patella rustica* y *Patella caerulea*²⁴⁰, *Pecten maximus* y *Pinna nobilis*.²⁴¹ Sin embargo la correlación es mucho más débil en el caso de *Tridacna gigas*,²⁴² *Concholetpas concholepas*,²⁴³ *Protothaca staminea*²⁴⁴ *Cassostrea virginica*.²⁴⁵ Prácticamente no se ha encontrado correlación en el caso de *Mesodesma donacium*,²⁴⁶ *Chione subrugosa*,²⁴⁶ *Mytilus californianus*,²⁴⁷ *Artica Islandica*²⁴⁸⁻²⁵⁰ y *Protothaca thaca*.²⁵¹

En este estudio se ha aplicado, por primer vez, el uso de la técnica LIBS para la obtención de la relación Mg/Ca en las conchas de moluscos modernos, *Phorcus Lineatus*. LIBS presenta importantes ventajas frente a los métodos convencionales, como son su escaso o nulo tratamiento de muestra, análisis a presión atmosférica, una alta resolución lateral (rango de μm), posibilidad de realizar un análisis posterior en los mismos puntos de muestreo que LIBS con otros métodos analíticos y una buena sensibilidad para la mayoría de elementos.^{44, 252} En la literatura se pueden encontrar trabajos donde se han aplicado otras técnicas analíticas como ICP óptico,^{241,253-256} ICP acoplado a masas,^{240, 250, 253} y ablación láser ICP-MS,^{242-244,246,249-251,257} para obtener dicha relación. Estas técnicas ofrecen una alta sensibilidad y bajos límites de detección. Sin embargo, a excepción de LA-ICP-MS todas necesitan etapas de preparación de muestra, que necesariamente llevan asociados procesos de química húmeda, necesitando una mayor cantidad de muestra y generando un aumento en los tiempos de análisis. En el caso de LA-ICP-MS no se necesita preparación de muestra, pero el tamaño de la cámara de ablación es limitado y no permite la manipulación de muestras de un cierto tamaño.



Además los tiempos de análisis y costes son muy superiores a los necesarios en el caso de LIBS.

En este trabajo, con el objetivo de evaluar el potencial de la relación Mg/Ca para inferir la paleotemperatura del agua marina en conchas de *Phorcus Lineatus*, se han realizado dos análisis. Por una parte se ha analizado la relación Mg/Ca en el borde, última etapa de crecimiento del molusco, de diferentes conchas recogidas mensualmente en la playa de Langre durante un año completo, con el fin de observar cambios composicionales y relacionarlos con la temperatura superficial del agua del mar obtenida instrumentalmente. Por otra parte, se ha obtenido la relación Mg/Ca durante una secuencia larga en tres conchas (análisis secuencial) con el objetivo de observar patrones cíclicos en la composición que pudieran estar relacionados con cambios estacionales.

En un estudio previo, se realizaron sobre las mismas muestras análisis de isótopos estables oxígeno para evaluar el uso del $\delta^{18}\text{O}$ como paleotermómetro.²³² Se observó como la capa aragonítica de *P. lineatus* crece en condiciones (o muy cerca) de equilibrio isotópico de oxígeno y que la temperatura del agua marina se puede inferir de los valores de $\delta^{18}\text{O}$ cuando se conoce la composición isotópica de oxígeno del agua del mar. Estos resultados se incorporan en este trabajo a modo de comparación y calibración con los datos de Mg/Ca obtenidos, sin embargo, no se dará mayor información del análisis isotópico ya que queda fuera del objetivo de este tesis.

7.1.2 Materiales y métodos

Muestras analizadas

Se analizaron muestras modernas de la especie *P. Lineatus* recogidas periódicamente entre Octubre de 2011 y Octubre de 2012 en la playa de Langre en la

costa cantábrica. (Fig. 7.1.1). Esta especie de molusco puede llegar a vivir hasta 10 años y tiene un tamaño medio de unos 35 mm de longitud con una tasa de crecimiento anual del orden de 15 a 26 mm (medido siguiendo la línea de la espiral generada),^{258,259} lo cual permite el análisis mediante LIBS durante al menos un año siguiendo su línea de crecimiento. Se analizaron un total de 47 conchas. La Tabla 7.1.1 muestra la identificación de cada muestra, con su fecha de recogida y la temperatura instrumental proporcionada por el instituto oceanográfico de Santander.

Antes de las medidas, se realizó una etapa previa de limpieza de muestra. Se sumergieron en una disolución de 30 % H_2O_2 y 70 % agua desionizada durante 40 horas y se dejaron secar a temperatura ambiente, posteriormente se introdujeron en un baño de ultrasónidos para eliminar cualquier resto biológico o de arena durante 5 minutos y se dejaron secar de nuevo a temperatura ambiente.

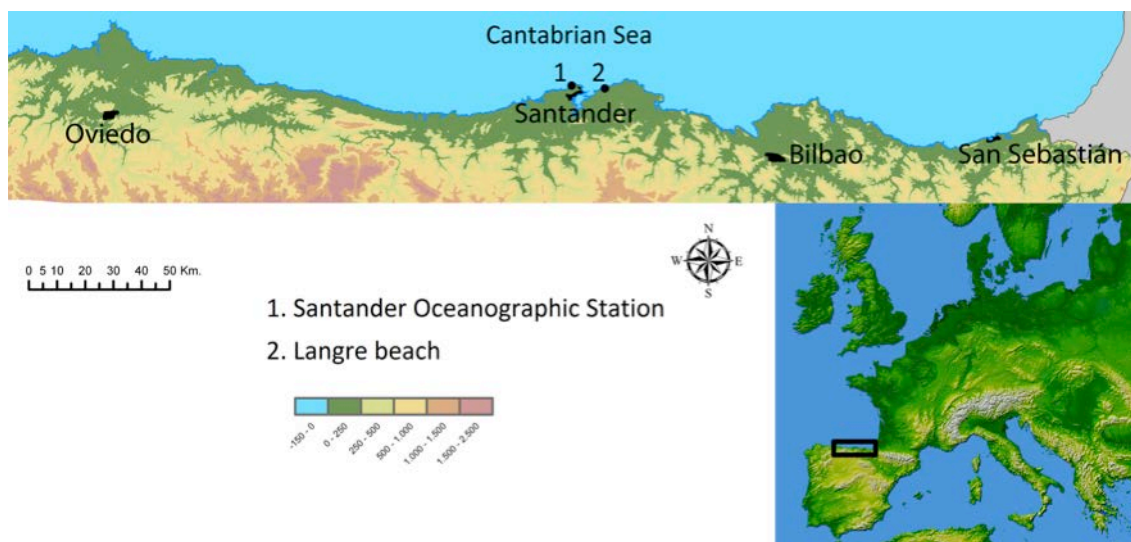


Figura.7.1.1: Mapa de la costa cantábrica donde se muestra la localización de la playa de Langre y la estación de medida del instituto oceanográfico.

La microestructura de la especie *P. Lineatus* ha sido previamente estudiada a través de microscopía electrónica de barrido (SEM) y mediante difracción de Rayos-X (XRD), mostrando una estructura en capas. Dicha estructura está formada por una capa externa de calcita externa y otra interna de aragonito. En este trabajo, se estudió la morfología del molusco con un microscopio petrográfico donde observamos varias capas: un recubrimiento amorfo o periostraco formado por conquiolina, que es una proteína encargada del almacenamiento de desechos del animal, una segunda capa de calcita prismática, calcita foliada y por último la capa más interna de aragonito, característica por su brillo nacarado (Fig. 7.1.2).

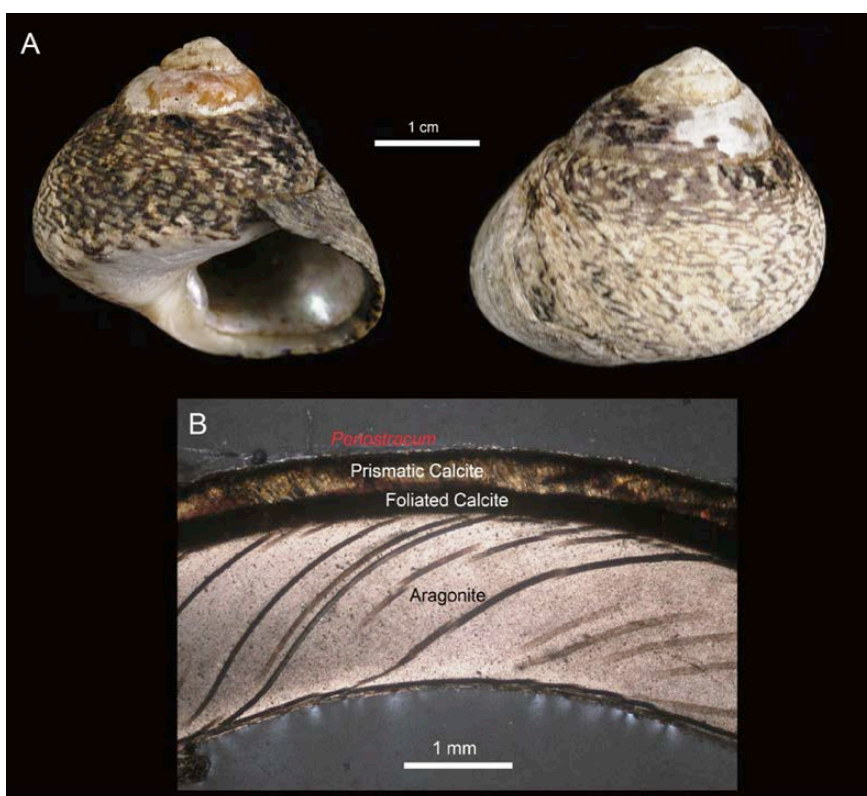


Figura. 7.1.2: a) Morfología de la concha de *P. lineatus* b) Sección transversal del gasterópodo mostrando las diferentes estructuras cristalinas: aragonit, calcita foliada, calcita prismática, periostraco orgánico.

Sistema experimental

El sistema experimental utilizado en este estudio se describió con anterioridad. En este caso para evitar la primera emisión del plasma, se utilizó un retraso de 5 μ s entre la obtención del espectro y el pulso del láser. El retraso, tiempo de integración y la fluencia fueron optimizadas en función de la relación señal/ruido.

Análisis LIBS

Antes de la realización de las medidas LIBS, se eliminaron las capas exteriores (periostraco y la capa de calcita) usando un microtaladro Dremel con una broca de 2 mm. Se midieron dos carbonatos biogénicos presentes en el molusco: calcita prismática y aragonito. Se evaluó la desviación estándar relativa (RSD) usando diferente número de promedios de medida para la obtención del espectro. Teniendo en cuenta el tiempo de análisis y la destrucción de muestra, finalmente se fijó usó un promedio de 20 pulsos en la misma posición, llegando a un valor de RSD inferior al 5 %. Las posiciones de muestreo se separaron 200 μ m cubriendo al menos 3.2 cm de la concha, siguiendo la línea de crecimiento, midiendo un total de 160 posiciones de muestreo. Las líneas de emisión utilizadas para realizar la relación fueron para Mg (II) a 279.55 nm y Ca (I) a 422.67 nm debido a su alta relación señal/ruido y evitando la saturación, interferencias con otras líneas de emisión.

La [Fig. 7.1.3](#) muestra la traza de cráteres formados por la interacción con el láser en la capa de calcita (A) y aragonito (B) a lo largo de la línea de crecimiento.

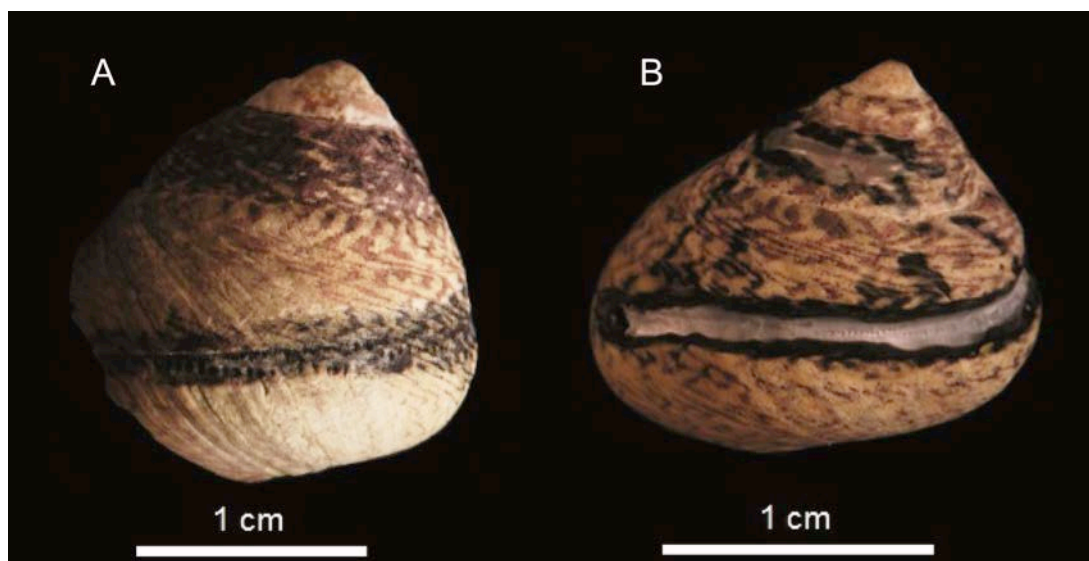


Figura. 7.1.3: Cráteres formados por los pulsos láser en a) *calcita* y b) *capa de aragonito*

7.1.3 Resultados y discusión

La Fig. 7.1.4 muestra un espectro LIBS típico de la capa de aragonito, cubriendo un rango espectral entre 240.0 nm y 500.0 nm. La asignación de las líneas de emisión de los elementos principales a sus correspondientes longitudes de onda se llevó a cabo usando la base de datos NIST.⁶⁸ Se observaron líneas de Mg (I) 285.2, Mg (II) 279.5, Ca (I) 300.7; 422.7 y Ca (II) 317.9; 393.4; 396.8 todo ello en unidades de nm. Debido a la morfología de la muestra (forma espiral) para mantener las condiciones de enfoque y recolección de la emisión, cada posición fue corregida usando un posicionador XYZ micrométrico evitando efectos de desenfoque y mejorando la reproducibilidad de las medidas.

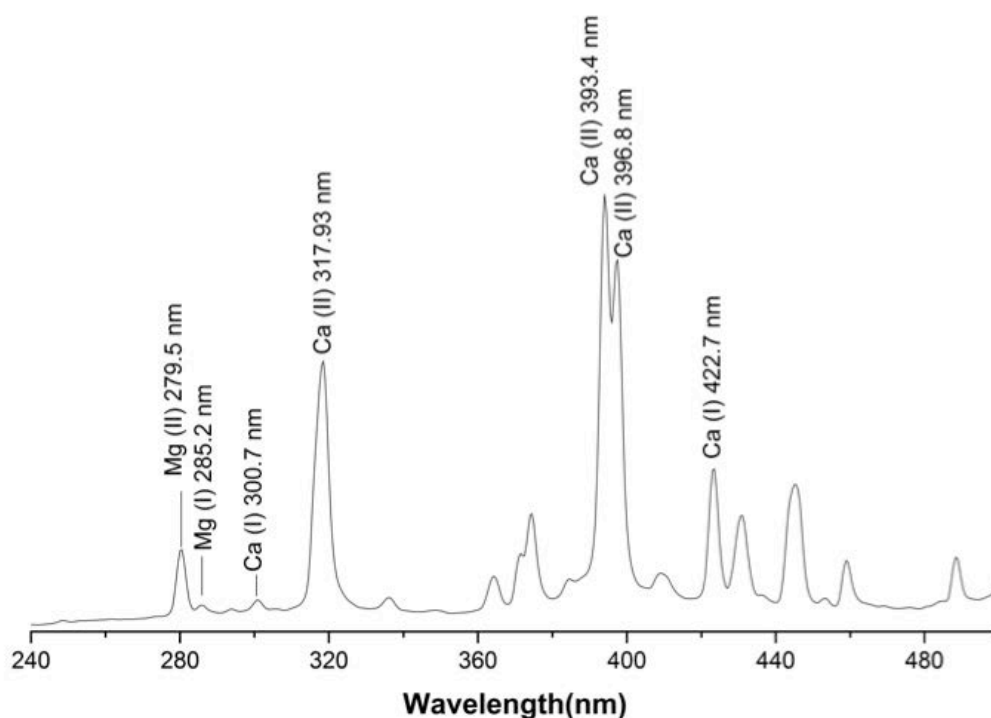


Figura 7.1.4: Espectro LIBS típico de la muestra de *P. Lineatus*

Medidas de la relación Mg/Ca en el borde

Las medidas de la relación Mg/Ca en el borde (último crecimiento de la concha antes de ser recolectas) permiten determinar los patrones de estacionalidad de las conchas, lo cual es primordial para establecer los patrones de recolección de moluscos y su importancia en las dietas de los primeros cazadores-recolectores. Para ello, se midieron 47 muestras diferentes de *P. Lineatus*, la correlación entre la temperatura superficial del mar y la relación Mg/Ca se muestran en la [Fig. 7.1.5 A](#) y [Fig. 7.1.5 B](#). Cada valor mostrado fue obtenido promediando los primeros cuatro valor de tres conchas diferentes recogidas el mismo día. Las barras de error muestran la desviación estándar relativa entre los tres especímenes. El mayor valor encontrado en la capa de calcita fue de 0.319, mientras que en aragonito fue tan solo de 0.06. Esta diferencia es

atribuida al hecho de que la calcita y el carbonato de magnesio presentan la misma estructura cristalina hexagonal y por tanto las soluciones sólidas entre ambas fases pueden ocurrir de forma más favorable. Además, el radio atómico del Ca y Mg es similar y el Mg puede ser fácilmente incorporado a la estructura de la calcita. Por otra parte, la capa de aragonito tiene una estructura cristalina ortorrómbica, haciendo que la incorporación de Mg sea desfavorable. Se ha observado como la relación entre el valor de Mg/Ca y la SST es inversa en la capa de aragonito (Fig. 7.1.5 A), mientras que en la capa de calcita la relación es directamente proporcional (Fig. 7.1.5 B) Por lo tanto, un incremento en la relación Mg/Ca es observado con aumentos de la temperatura cuando se trata de la capa de calcita y en sentido contrario para la capa de aragonito, a menores temperaturas, mayores valores de Mg/Ca. El valor de correlación encontrado entre la relación Mg/Ca y SST fue de 0.43 y 0.44 en la capa de aragonito y calcita, respectivamente. Esta correlación demuestra que la incorporación de Mg a la estructura del molusco no solo está influenciada por la temperatura sino que también depende de factores como la salinidad, calidad y cantidad de alimentación y de los ciclos reproductivos, entre otros. Sin embargo, la concentración de los elementos en una posición específica es homogénea dentro de una misma capa, independientemente de la profundidad del cráter. Es importante mencionar que se encontraron altos valores de RSD del 30.7 % y 25.9 % en la capa de aragonito y calcita respectivamente para diferentes especímenes colectados el mismo día, lo que indica la alta variabilidad entre diferentes animales de la misma especie y bajo las mismas condiciones ambientales. Este hecho indica que la incorporación de los elementos a las conchas de los moluscos está fisiológicamente controlado por factores metabólicos, efectos cinéticos, tasas de calcificación y diferencias entre la orientación y morfología de los cristales.^{253, 254,256,260,261}

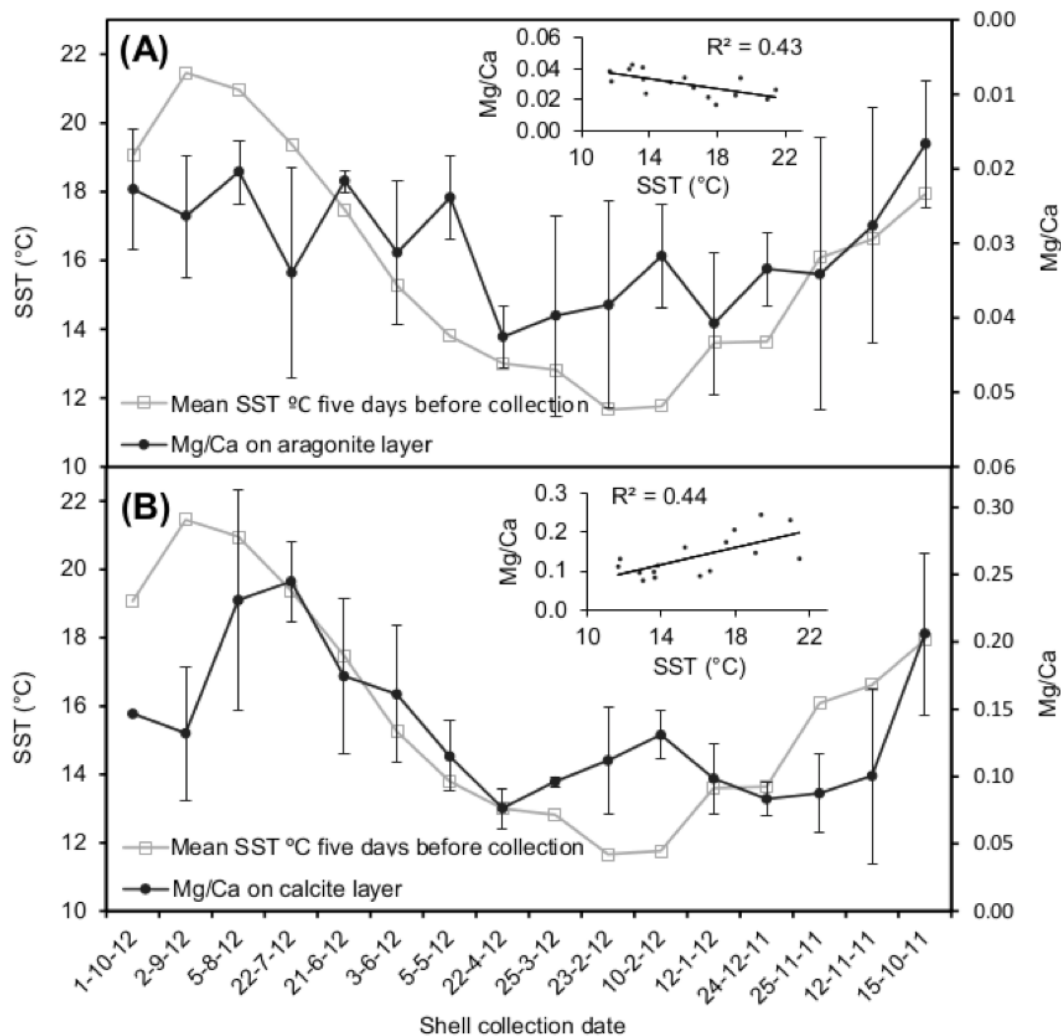


Figura 7.1.5: Correlación anual entre SST y la relación Mg/Ca en A) capa de aragonito y B) capa de calcita

Medidas de la relación Mg/Ca secuenciales

Se analizaron tres conchas, considerando al menos un año, a lo largo de su línea de crecimiento. La relación Mg/Ca de las muestras LANO-4 y LANO-5 fueron medidas solo en la capa de aragonito, mientras que la muestra LANO-54 fue medida en ambas capas aragonito y calcita.

197

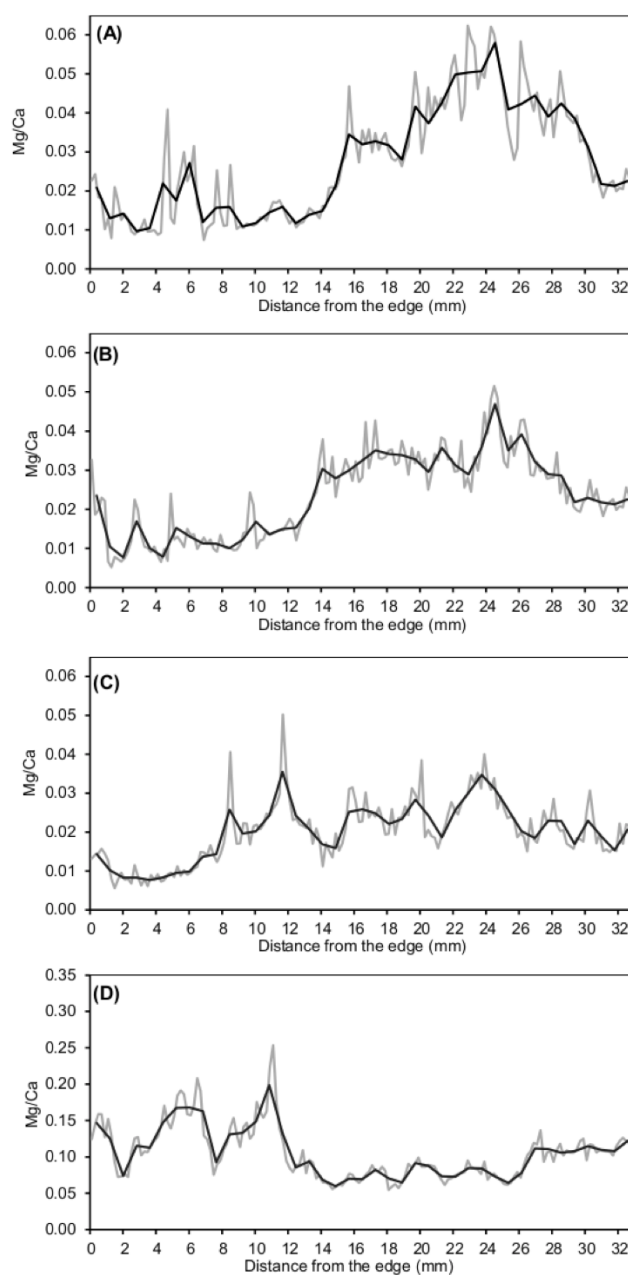


Figura 7.1.6: Relación Mg/Ca a lo largo de la dirección de crecimiento de cuatro conchas de *P. Lineatus*. Las dos primeras A) LANO-4, B) LANO-5 fueron medidas en la capa de aragonito. C) muestra la medida en la capa de aragonito de LANO-52 y D) LANO-52 en calcita. La línea gris indica cada punto obtenido y la negra el promedio de 4 valores Mg/Ca.



Finalmente se realizaron secuencias de isótopos estables de oxígeno para las muestras LANO-5 y LANO-52 en la capa de aragonito y se comparó con la relación Mg/Ca obtenida por la técnica LIBS. Es importante destacar la resolución lateral obtenida por LIBS comparada con las medidas de isótopos. Aproximadamente, cada medida de isótopos comprende 4 posiciones de muestro de LIBS, lo que en algunas situaciones se puede traducir en una pérdida de información pudiendo ocultar tendencias estacionales. Se encontraron valores de correlación entre las medidas LIBS e isótopos de $R^2=0.7$ en el caso de LANO-5 (Fig. 7.1.7 A). En el caso de LANO-52, la correlación encontrada fue de $R^2=0.07$, considerando la secuencia completa. Sin embargo, cuando se consideran los últimos 16 mm de crecimiento esta correlación aumenta considerablemente hasta un valor de $R^2=0.65$ (Fig. 7.1.7 B). Esta anomalía es difícil de explicar, nuestra hipótesis es que existen otros factores involucrados en la incorporación de Mg además de la temperatura. Si consideramos la correlación obtenida para LANO-5 y la última parte de LANO-52, la correlación encontrada este trabajo está entre las más altas encontradas en bibliografía, solo Ferguson et al²⁵⁰ y Freitas et al²⁴¹ han publicado valores de correlación superiores en lapas de $R^2=0.79$ y vieiras $R^2=0.77$, respectivamente.

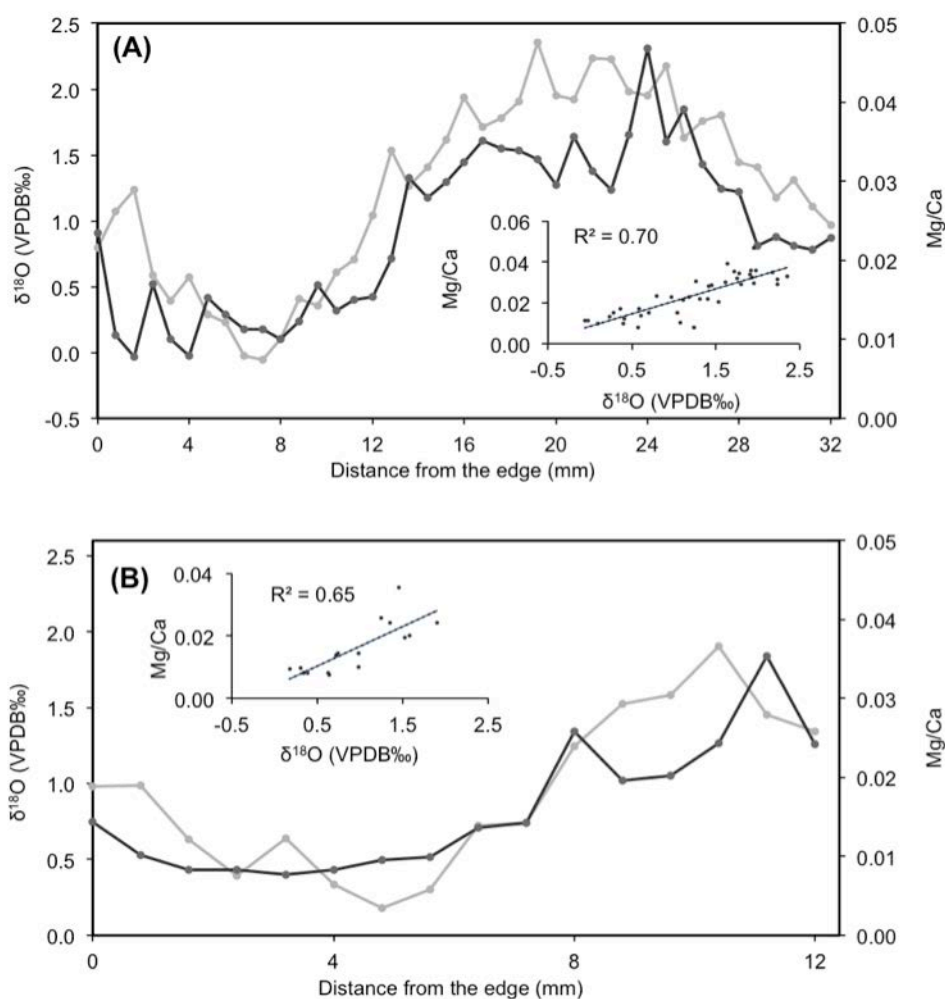


Figura 7.1.7: Relación entre $\delta^{18}\text{O}$ y Mg/Ca a lo largo de la línea de crecimiento para A) LANO-5 y B) LANO-52

7.1.4 Conclusiones al análisis de muestras de molusco mediante LIBS

Se ha evaluado, por primera vez, el uso de la relación Mg/Ca obtenida mediante LIBS para la obtención de información geoquímica sobre un proxy paleoclimático, *P. Lineatus*. Los análisis de la relación Mg/Ca obtenida por LIBS ofrece resultados satisfactorios, mostrando su potencial como técnica analítica para el estudio medioambiental.

Conocer las condiciones climáticas del pasado es crucial para muchas disciplinas, como la paleo-climatología, arqueología, paleo-oceanografía o la paleontología. Establecer la

temporada de captura de los moluscos tiene importantes implicaciones para estudios arqueológicos, como la reconstrucción de patrones de asentamiento y estrategias de subsistencia de cazadores-pescadores-recolectores prehistóricos.

Los coeficientes de correlación entre la relación Mg/Ca y los isótopos de oxígeno, nos indica que la incorporación de elementos traza, como el Mg en el carbonato cálcico biogénico de la especie *P. lineatus* no depende sólo de la temperatura del agua del mar, sino que también factores tales como la salinidad, la calidad y cantidad de los alimentos, entre otros.

LIBS es una técnica analítica apropiada para llevar a cabo análisis elementales de conchas de moluscos, ofreciendo una metodología analítica alternativa para la determinación de los patrones climáticos, produciendo resultados comparables con otras técnicas contrastadas como son los análisis isotópicos de oxígeno. Además, LIBS permite una considerable reducción tanto temporal como económica, con una pérdida mínima de material arqueológico. Sin embargo, es necesario continuar el estudio con el fin de mejorar los coeficientes de correlación y conocer la influencia de otros parámetros en la incorporación de Mg a las diferentes formas alotrópicas de carbonato de las conchas de moluscos.

Agradecimientos:

Especial agradecimientos a todos los autores del artículo, A. García-Escárzaga, I. Gutiérrez-Zugasti, M. R. González- Morales, J. Martín-Chivelet and J. O. Cáceres. Los análisis de isótopos estables de oxígeno han sido financiados por el proyecto del Plan Nacional de I+D+I “La respuesta humana al cambio climático global en una zona litoral: el caso del tránsito al Holoceno en la costa cantábrica (10000-5000 cal BC) (HAR2010-22115-C02-01)” otorgado por el Ministerio de Economía y Competitividad.

El Servicio de Actividades Pesqueras del Gobierno de Cantabria concedió los permisos necesarios para la recogida de muestras modernas y la Planta de Acuicultura del Centro Oceanográfico de Santander perteneciente al Instituto Español de Oceanografía cedió amablemente los datos térmicos y de salinidad. El material y los recursos técnicos necesarios fueron proporcionados por el Instituto Internacional de Investigaciones Prehistóricas de Cantabria y por la Universidad Complutense de Madrid.

7.2 Resultados preliminares del estudio de relaciones elementales por LIBS: Análisis de Corales

Resumen

Debido a que este trabajo se continúa realizando en nuestro laboratorio en colaboración con el Dr. Ismael Coronado del Departamento de Paleontología de la Universidad Complutense, los resultados mostrados aquí son preliminares. Se ha aplicado la técnica LIBS al análisis de dos corales con importancia desde el punto de vista paleoclimatológico recogidos en la costa mediterránea y atlántica. En primer lugar, se ha obtenido la relación Mg/Ca, Sr/Ca e H/Ca a lo largo de la línea de crecimiento de los corales y se ha correlacionado con la temperatura superficial del agua del mar (*Superficial Seawater Temperature*, SST), obtenida instrumentalmente en estaciones oceánicas. Teniendo en cuenta la tasa de crecimiento y la datación de dichos corales, se ha podido observar la variación de temperatura desde los años 40 hasta el año 2010 en el caso del coral *Dendrophyllia cornígera*, registrando cambios en la temperatura de menos de 0.1°C. En el caso del coral *Cladocora caespitosa*, se ha medido la relación Mg/Ca e H/Ca y se ha comparado con la temperatura superficial del agua de mar en zonas próximas a las que se encontró el coral, correlacionado las relaciones elementales entre los años 1985 y 1997. Por último, se han incluido las primeras imágenes



multielementales obtenidas por espectroscopia láser sobre una muestra de coral. Los resultados obtenidos abren una nueva visión en el uso de LIBS para aplicaciones geológicas y su estudio paleoclimatológico. Esta nueva técnica ha permitido la obtención de imágenes (240.000 píxeles) multielementales de metales y elementos ligeros en un amplio rango de medida y con alta resolución lateral (hasta 15 μm). Con este trabajo se da un paso a delante mostrando una sensibilidad y resolución sin precedentes en el análisis geológico superando a técnica establecidas como *synchrotron radiation microanalysis* (SXRF), *laser ablation inductively coupled plasma mass spectrometry* (LA-ICP-MS) mucho más costosas y con complejos sistemas experimentales restringido a grandes centro de investigación y poco accesibles para análisis rutinarios.

7.2.1 Introducción al análisis de corales

Los corales son animales sésiles que habitan los fondos oceánicos someros y que son capaces de producir exoesqueletos minerales. Mediante la aglomeración de estos esqueletos tras generaciones en el fondo marino, así como de otros organismos que construyen sus esqueletos y conchas de carbonato cálcico, producen edificios marinos que se conocen como arrecifes. Los arrecifes son uno de los puntos de mayor diversidad de la vida en la Tierra, se concentran en mares someros tropicales y son una de las únicas estructuras biológicas observables desde el espacio. Existen otros corales que viven en áreas templadas e incluso frías, y en aguas moderadas y profundas, pero estos no forman arrecifes.²⁶⁶ Los arrecifes son órganos muy sensibles a los cambios climáticos y por esa razón se emplean en los estudios paleoclimatológicos. Se han llevado a cabo numerosos estudios empleando los esqueletos aragoníticos (CaCO_3) de los corales como archivos geológicos para la reconstrucción de la temperatura del agua marina



En este estudio se ha utilizado, por primera vez, la técnica LIBS para el seguimiento de las relaciones elementales en muestras de corales con el objetivo de analizar la correlación presente con la temperatura de superficial del agua de mar y evaluar los posibles factores que afectan a dicha correlación. En bibliografía solo existe un trabajo donde se ha aplicado la calibración *free LIBS* para la determinación cuantitativa de ciertos elementos en una matriz coralina llevado a cabo por S. Panhdhija et al.²⁷⁴ Sin embargo, el objetivo de dicho trabajo no está relacionado con aspectos paleoclimáticos



y por tanto nuestro trabajo es el primero en este ámbito. Además de la correlación de diferentes relaciones elementales con la temperatura del agua de mar, se ha incluido un estudio de imágenes multielementales obtenidas mediante espectroscopía láser. El uso de esta técnica no sólo es novedoso desde el punto de vista de la técnica, sino que también lo es desde su aplicación, en particular al análisis de corales.

Las técnicas de imagen existen desde hace más de tres siglos, la microscopía permite observar estructuras que el ojo humano desnudo no puede ver y por tanto descubrir un nuevo horizonte. Entre las técnicas microscópicas de alta resolución más aplicadas, cabe destacar *synchrotron radiation microanalysis* (SXRF), *nano-secondary ion mass spectrometry* (nano-SIMS), *laser ablation-inductively coupled plasma mass spectrometry* (LA-ICP-MS), *electron probe microanalysis* (EPMA). Aunque todas ellas presentan una alta sensibilidad y bajos límites de detección, su instrumentación y su uso es complejo, requiriendo de altos recursos económicos y de personal cualificado para llevar a cabo los análisis. El grupo del Prof. Vincent Motto-Ros y colaboradores de la universidad Claude Bernard de Lyon han desarrollado una técnica compatible con los microscopios ópticos actuales basada en la espectroscopia de ablación láser (LIBS) para la formación de imágenes elementales de forma rápida y con instrumentación accesible, permitiendo la detección de elementos traza con una sensibilidad del orden de pocas ppm y con una resolución lateral menor de 10 μm . Sus estudios se basan en el uso de esta tecnología en el ámbito de la biología y medicina como demuestra la bibliografía.^{275,276} En este caso, esta tecnología ha sido aplicada al análisis de diferentes elementos presentes en los carbonatos de los esqueletos de coral para obtener la distribución espacial de elementos como Ca, Mg, Sr y con ellos obtener información geológica y ambiental de la muestra de coral así como de la biología de estos animales.

7.2.2 Materiales y métodos

Muestras analizadas

Se han analizado dos colonias de coral diferentes, en primer lugar un coral de la especie *Dendrophyllia cornigera* (Fig. 7.2.1) y otra colonia de la especie *Cladocora caespitosa* (Fig. 7.2.2). Para su análisis, los corales se limpiaron con una solución de hipoclorito de sodio al 7 % durante 24 horas para eliminar los restos de tejidos orgánicos adheridos a la superficie y se cortaron en diferentes secciones para dejar visible la superficie interna del coral. Algunas colonias se cortaron y pulieron para observar las etapas de crecimiento de dichas colonias, realizándose láminas delgadas y secciones pulidas con las superficies analizar. Esta metodología la llevo a cabo el Dr. I. Coronado et al ²⁷⁷ dentro de su trabajo de tesis.

Análisis LIBS

El sistema experimental utilizado es el mismo que en el caso de moluscos. El tiempo de retraso entre la obtención del espectro y el pulso del láser fue optimizado a 5 μ s, para evitar la radiación de fondo y el tiempo de integración se mantuvo en 100ms. Para la obtención de cada espectro se promediaron 20 pulsos láser en la misma posición, siguiendo el procedimiento desarrollado con el análisis de moluscos. Las posiciones de muestreo se separaron 200 μ m, para evitar el análisis de posiciones previamente irradiadas. En el caso de la muestra *D. cornigera* se tomaron un total de 265 puntos de muestreo, mientras que para la muestra de *C. caespitosa* se midieron 3 coralitos con un promedio de 70 posiciones en cada uno de ellos. Las líneas de emisión utilizadas fueron: Mg (II) a 279.55nm, Ca(I) a 422.67nm, Sr(I) 407.15nm e H α 656nm, la selección se realizó en base a la relación señal/ruido, evitando líneas saturadas e interferencias con otras líneas de emisión.



Figura 7.2.1: *Morfología externa e interna del coral Dendrophyllia cornígera. Imágenes tomadas por el Dr. Ismael Coronado (UCM).*



Figura 7.2.2: *Morfología externa y corte de una colonia de coral Cladochonus. Imágenes tomadas por el Dr. I. Coronado*

7.2.3 Resultados y discusión de las relaciones elementales

Las medidas LIBS se realizaron en dos corales diferentes, un coral de la familia *Dendrophyllia cornígera* y una colonia de corales *Cladochonus*. La Fig.7.2.3 muestra un espectro LIBS típico de *Cladochonus*, en el rango espectral 200.0nm y 900.0nm. La asignación de las líneas de emisión de los elementos principales a sus correspondientes longitudes de onda se llevó a cabo usando la base de datos NIST.³³ Se observaron líneas de Mg (I) 285.2nm, Mg (II) 279.5nm, Ca (I) 300.7nm; 422.7nm y Ca (II) 317.9nm; 393.4nm; 396.8nm; bandas moleculares de CaOH, O₂ y Ca_xO_x en la región de 600nm.

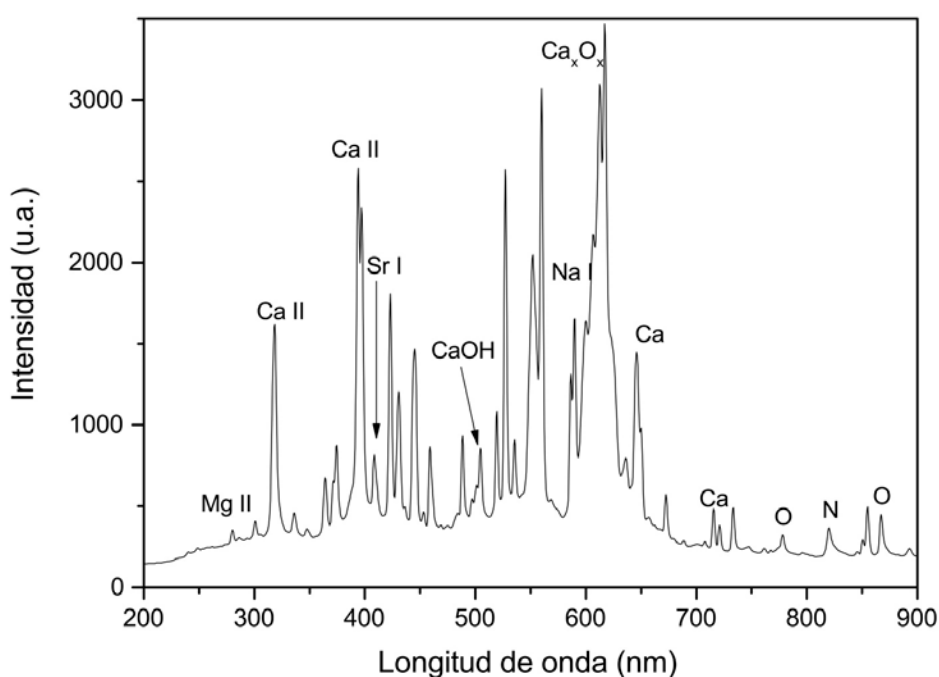


Figura 7.2.3: Espectro LIBS típico de una muestra de coral (*C. caespitosa*) compuesto fundamentalmente de Carbonato de Calcio.



En este caso el coral considerado presenta una morfología óptima para la medida LIBS, con una superficie plana y un tamaño considerable, facilitando la adquisición de espectros. Se midieron un total de 265 puntos, teniendo en cuenta la datación del coral podemos correlacionar las medidas LIBS con su periodo temporal. Se han cubierto 70 años de vida del coral, desde el año 1940 hasta el 2010. La Fig.7.2.4 muestra los valores obtenidos de la relación Sr/Ca a través de LIBS (azul) y la temperatura superficial marina instrumental (rojo). Para una mejor visualización de la tendencia los valores tanto de las relaciones como de la temperatura fueron suavizados, trazo negro y verdes, respectivamente. Se observa una buena correlación entre las medidas LIBS con la temperatura, es importante observar la escala de temperaturas y como la relación Sr/Ca es sensible a cambios de temperatura del orden de 0.1°C . Estos resultados son preliminares y muestran el potencial de la técnica LIBS para llevar a cabo medidas paleoclimáticas. El siguiente paso es la interpretación de los resultados, que se continuará después de la finalización de la tesis.

El segundo coral elegido presenta una morfología más compleja, se trata de una colonia de coralitos de pequeño tamaño con numerosas ramificaciones y direcciones de crecimiento. En este caso la muestra también tiene una superficie plana, lo que facilita el análisis LIBS. Se midieron 3 coralitos diferentes de la misma colonia para observar posibles variaciones entre coralitos. Los resultados aquí mostrados son los obtenidos con las medidas del último coralito (mayor tamaño), donde se obtuvieron 76 espectros LIBS. La [Fig. 7.2.5](#) muestra la relación Mg/Ca (trazo azul) frente a los valor de temperatura superficial del mar (trazo rojo). Se han considerado al menos 12 años de

vida desde el año 1985 hasta 1997. La correlación es buena en general, la zona central entre los años 1987 y 1992. Los valores iniciales y finales presentan una menor correlación, el motivo no es aún claro. Puede ser debido a la dificultad de obtener las medidas en dichas zonas o variaciones en la composición del esqueleto coralino, debido a factores no relacionados directamente con la temperatura, sino fisiológicos. Se ha estudiado por primera vez la relación H_v/Ca en muestras de coral. Los resultados se muestran en la Fig. 7.2.6, en este caso la correlación mejora bastante, incluso a principios de los 90, donde la relación Mg/Ca no era clara. La presencia de H en estas muestras puede ser debida a números factores, la hipótesis que barajamos es la presencia de pequeñas burbujas de agua ocluidas en la estructura biológica cristalina. En estos momentos no podemos confirmar dicha hipótesis y se realizará un estudio detallado para poder refutarla o rechazarla.

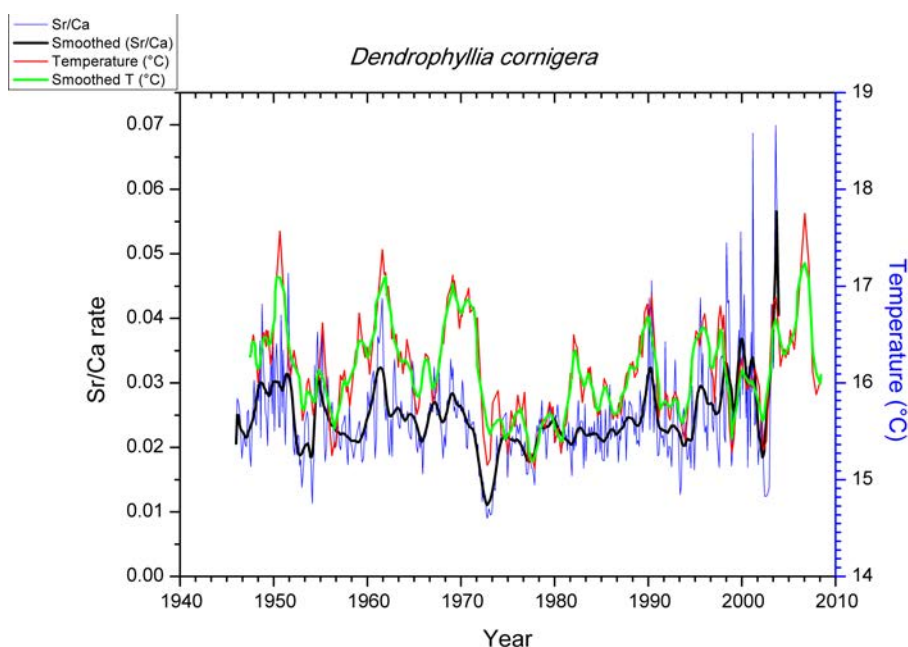


Figura 7.2.4: Evolución y correlación en la muestra de coral *Dendrophyllia Cornigera* usando la relación Sr/Ca . Se observa una buena correlación, permitiendo observar cambios en la composición incluso con variaciones de $0.1^{\circ}C$.

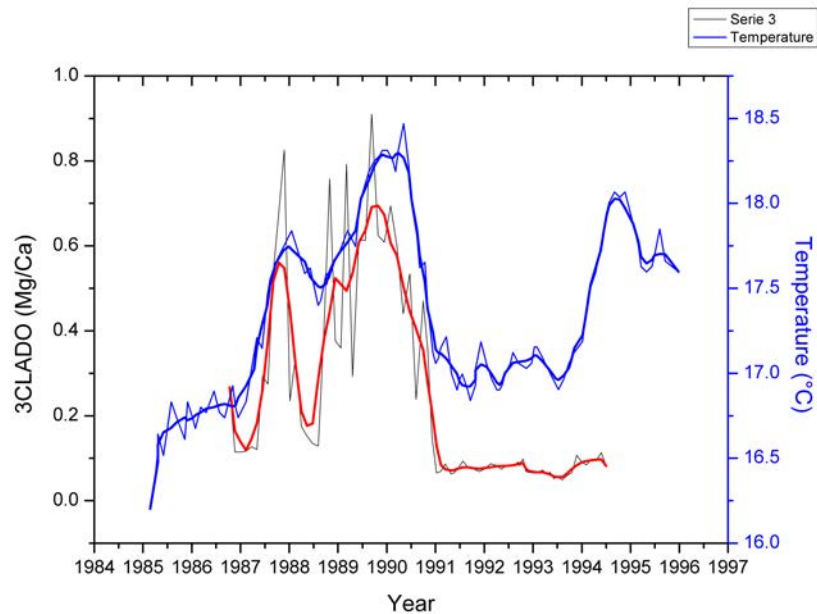


Figura 7.2.5: Evolución y correlación en la muestra del coralito número 3 de *Cladocora caespitosa* usando la relación Mg/Ca (rojo) cubriendo más de diez años de crecimiento.

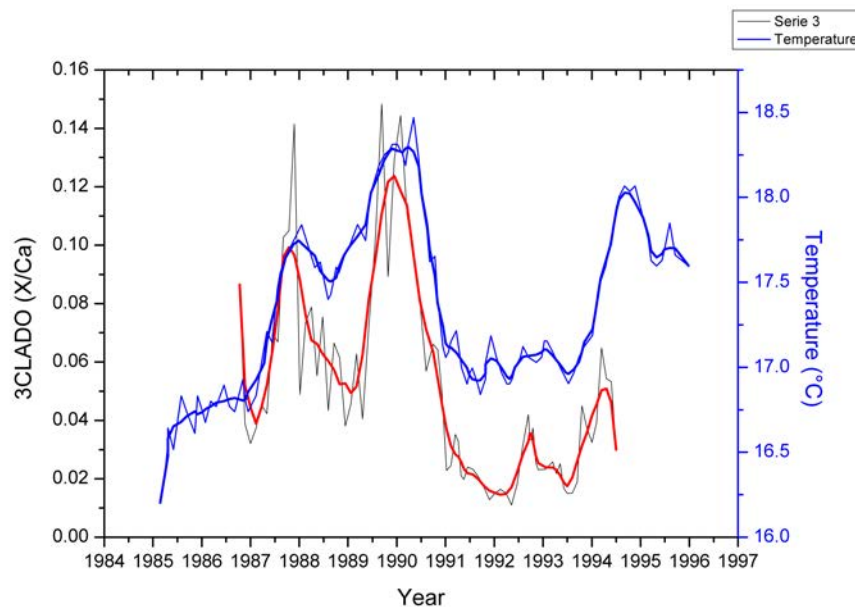


Figura 7.2.6: Evolución y correlación en la muestra del coralito *Cladocora caespitosa* usando la relación H_{α}/Ca (rojo) y la temperatura (azul). Observa una alta correlación, la resolución temporal de LIBS es superior a la disponible para la temperatura.

7.2.4 Metodología para la obtención de imágenes multielementales

El estudio se llevo a cabo en el L'Institut Lumière Matière de Lyon gracias al Prof. Vincent Motto-Ros. Como sistema experimental se usó un láser Nd:YAG trabajando en el IR (1064 nm) con una frecuencia de trabajo de 10 Hz. El haz láser se focalizó sobre la muestra con un objetivo de microscopio 15X y la emisión del plasma fue colectada a través de una fibra óptica y dirigida dos espectrómetros Czeny Turner Shamrock 500i y Shamrock 303i ambos dotados de dos cámaras ICCD de la marca Andor. Todas las medidas se llevaron a cabo bajo un flujo constante de Ar para mejorar la relación señal ruido. La muestra se posiciono es un sistema XYZ motorizado y controlado por ordenador para obtener un solo espectro en cada posición de muestreo.

7.2.5 Resultados de las imágenes multielementales

Después de la medida 2D, se generaron las imágenes elementales para cada elemento de interés mediante la selección de las líneas de emisión específicas. Cada elemento es asignado con un color arbitrario pudiendo visualizar la abundancia de cada elemento a través de las imágenes.

En este caso se analizaron los elementos más frecuentes en muestras de carbonato biológicos como Ca, Mg, Sr. La [Fig. 7.2.7](#) muestra las imágenes elementales para cada elemento analizado siguiendo una escala de colores. Cada imagen es generada muestreando un rectángulo de 600 x 400 píxeles con una resolución lateral de 30 μ m, el área analizada total fue de 216 mm² obteniéndose un total de 240.000 pulsos láser. El tiempo necesario para realizar las medidas fue de aproximadamente 7 horas teniendo en cuenta la frecuencia del láser. Debido a la resolución y tamaño de la imagen generada, es posible observar la distribución espacial de dichos elementos en el esqueleto de coral como si se tratase de una imagen óptica. La resolución lateral de la técnica es

directamente proporcional al tamaño del cráter generado y por tanto de la energía del pulso utilizada. Para mejorar la resolución lateral es necesario trabajar con energías inferiores a 10 mJ por pulso, sin embargo una disminución de la energía produce una menor cantidad de material ablacionado y con ello un plasma más pequeño y difícil de detectar, disminuyendo la sensibilidad de la detección. Es importante encontrar un balance entre la resolución y la capacidad de detección de la técnica. En nuestro caso cada pixel (espectro) de la imagen fue obtenido con un único pulso láser de tan sólo 3.5 mJ de energía, demostrando la alta sensibilidad del sistema utilizada. El tamaño del cráter generado es de aproximadamente $8\text{ }\mu\text{m}$ y la masa ablacionada es aproximadamente 0.4 ng. La energía del pulso láser fue monitorizada durante todo el proceso con el objetivo de minimizar las variaciones pulso a pulso.

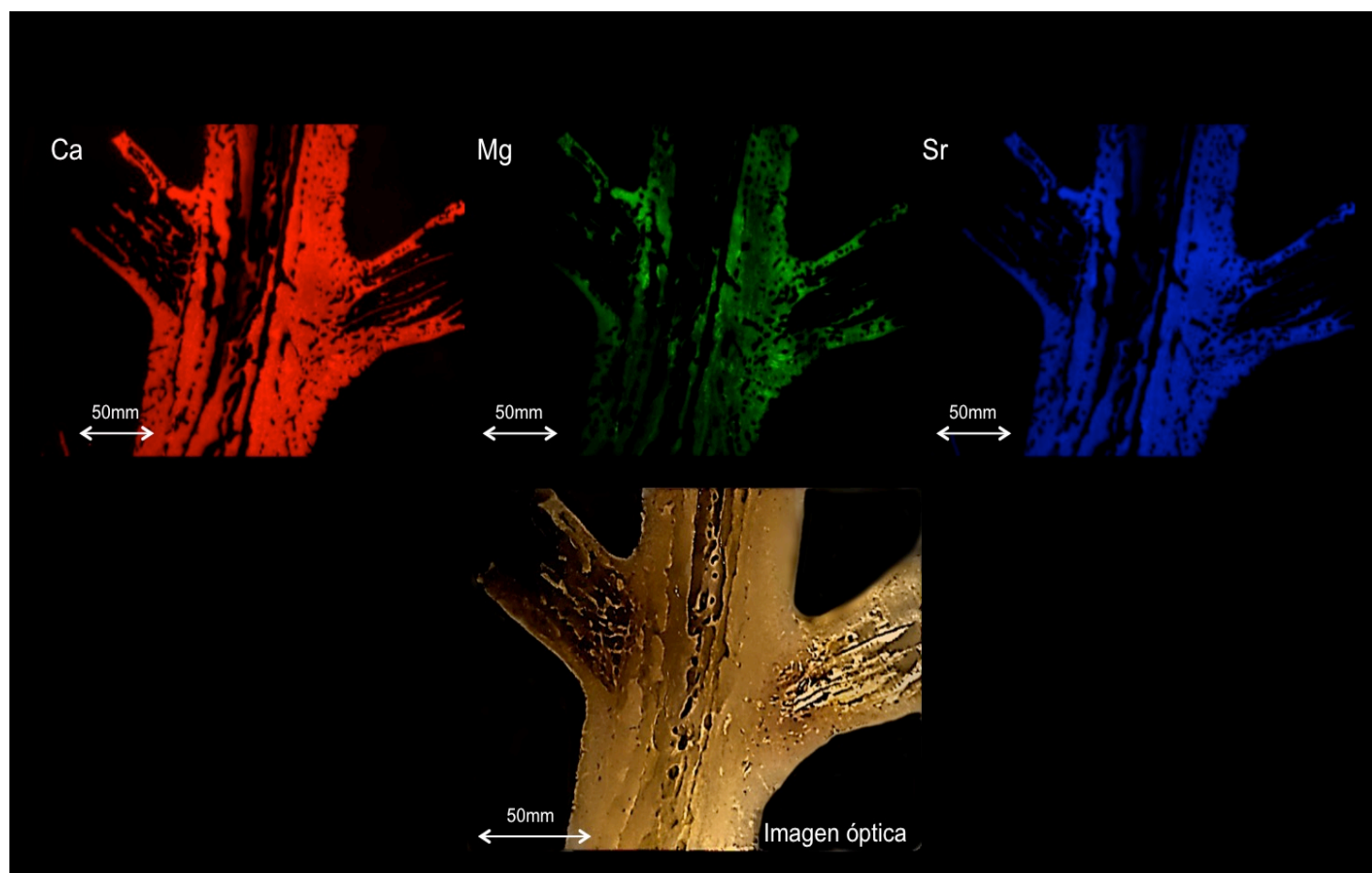


Figura 7.2.7: Imágenes elementales de una muestra de coral obtenida mediante LIBS.

7.2.6 Conclusiones al análisis preliminar de corales

En este estudio preliminar se ha evaluado la capacidad de la técnica LIBS para producir información útil de la composición de corales con un importe interés paleoclimatológico. Se ha observado como las relaciones elementales entre varios elementos como son Sr, Mg e H_{α} frente a Ca siguen la misma tendencia que las variaciones de temperatura del agua superficial del mar. Se ha demostrado la capacidad de LIBS para trabajar con corales grandes como *Dendrophyllia Cornigera*, así como con colonias con coralitos pequeños y de difícil acceso para otras técnicas analíticas.

Es necesario el estudio de más corales para conocer mejor estos proxies, pero los resultados obtenidos son prometedores y muestran el potencial de la técnica LIBS para evaluar relaciones elementales que proporcionan información de alta resolución, precisa, de forma rápida y económica.

En cuanto a las imágenes elementales, el uso de esta tecnología permite la obtener la distribución espacial varios elementos con una sensibilidad adecuada, alta resolución lateral y relativamente rápido comparado con las técnicas de imagen elemental convencionales. La distribución de los diferentes elementos dentro del esqueleto coralino no solo va a producir información climática, si no que además permite un mejor conocimiento de la biología de este animal. Actualmente se continua trabajando en la discusión e interpretación de los resultados, así como con la preparación de nuevas medidas y estudios.

Agradecimientos:

Quiero agradecer la colaboración de Ismael Coronado y Sergio Rodríguez del Departamento de Paleontología de la Facultad de Ciencias Geológicas de la UCM por su trabajo y tiempo dedicado a este estudio. Algunas muestras de corales fueron



recolectadas por el Proyecto Fauna II y han sido prestadas al Dr. Ismael Coronado por parte del Museo de Ciencias Naturales de Madrid, así como por la colección del Departamento de Paleontología de la UCM. Al Prof. Vincent Motto-ros y Florian Trichard por facilitarnos su sistema experimental en las medidas de las imágenes elementales y el tiempo dedicado. También a mi director Jorge Cáceres por la posibilidad de realizar las medidas experimentales y su interpretación de los resultados.



Conclusions

According to the initial objectives, the obtained results and the individual conclusions presented in each chapter of this thesis, we can draw the following conclusions:

- A methodology combining laser induced breakdown spectroscopy with advanced chemometric methods for the classification and discrimination of samples in the field of the food industry has been developed and applied, for the first time. The methodology has been successfully validated in the control of the adulteration of olive oil and dairy products as well as in the determination the geographical origin of red wines. Given the speed and the accuracy in the results of the measurement obtained, this methodology can be useful for performing automated, reliable and robust analysis, even in real time, in several areas related to the food industry such as the quality control and food safety. It has also been shown that the gelation of wine samples produces better results than performing direct analysis . This sample preparation method is presented as an alternative way to analyze liquid improving the working conditions of the technique. The selection of the input variables for the mathematical models incorporated features of the samples allowing an/their unequivocal classification.
- Chemometric methods have been employed and evaluated for a quick and robust discrimination of human biological remains based on the spectral characteristics of bones and teeth. This methodology could be considered as an alternative, simple and cost-effective technique that could complement the current methodologies (slower and more expensive) in the task of reconstruction and



discrimination of human remains, especially in the case of victims of natural disasters and acts of genocide.

- The performance of seven supervised chemometric methods (SIMCA, PLS-DA, LDA, CART, BLR, SVM and NN) has been evaluated in the classification of skeletal remains based on LIBS measurements. It has been shown that the neural networks models, specifically designed to reach the objectives of this thesis, outperformed in terms of sensitivity, generalization ability and robustness providing an efficiency of 100%.
- The Mg/Ca ratio obtained by LIBS has been measured on a paleogeochemical proxy i.e. *P. lineatus* to reconstruct paleotemperatures and season of capture of molluscs. The potential of the LIBS has been shown as an analytical technique for environmental research in the determination of weather patterns. The results obtained by LIBS were comparable results with a known reliable measurement i.e. oxygen isotope ratio. LIBS technique presents significant advantages such as the reduction in the time of analysis and financial costs, and more importantly a minimal loss of archaeological material. It has also been observed that other elemental ratios such as Sr, Mg and alpha H over Ca are highly dependent on the sea surface temperature (SST) in corals *Dendrophyllia cornigera* and *Cladochonus*. It has been possible to determine variations of less than 0.1°C in the SST, showing the capabilities of the methodology and starting a new way in the development of new applications.
- Preliminary results of elemental images on corals have also been shown. These results are on the cutting edge of the technique, being able to detect



simultaneously metal, light elements and molecules with a high sensitivity in single shot configuration reaching a lateral resolution of 15 microns over long distances (4 cm).

It has been shown that the LIBS technique can be used in several applications being able to co-work in multidisciplinary studies of samples in different states of aggregation, with simple experimental procedures, reducing significantly the time and cost of the current standard analysis. Thanks to its simplicity of implementation and low cost, LIBS technique offers many possibilities and applications. New advances in instrumentation and experimental set-up have reduced or even removed the inherent drawbacks of the technique, such as the low sensitivity, meager reproducibility and matrix effect, reaching detection limits below of parts per million and relative standard deviations less than 1%.

Chemometric methods are an important tool for the analysis of data obtained by LIBS for both qualitative and quantitative analysis. The selection of appropriate analytical methods is crucial to obtain reliable and accurate results. Moreover, it is also important to record spectra in a systematic and reproducible procedure. Nonlinear classification methods have shown better performance, especially neural networks, because of the nonlinear effects present in all of the LIBS spectra. Neural networks provide the best results due to their ability to model nonlinear relationships between variables, being able to extract meaningful information from the LIBS spectra without the loss of information and with a simple implementation. The selection of the input variables is one of the key steps of the training process and must be carefully selected in order to obtain reliable results.



Nowadays, there are only few industries that incorporate LIBS system for routine analysis, however, due to the advances that have been made in the last 10 years, the technique is becoming more and more important and in the near future will be a technique widely extended not only from the research point of view but also for routine analysis. LIBS provides a huge amount of data in short period of time, so that the combination between LIBS and chemometric methods will be of vital importance in the interpretation of the results in order to obtain meaningful information in a quick and easy way.

In the case of elemental images, the use of chemometric analysis methods allows to obtain information easily among a large number of spectra (above 200.000) by personnel without a high analytical background or spectroscopic expertise, providing a semi-automatic analysis through chemometric models.

This thesis has allowed me to acquire the knowledge and the expertise necessary to perform studies on laser ablation spectroscopy, chemometrics methods and spectroscopic. Furthermore, other aspects of different disciplines such as wine production, tissue degradation and postmortem mechanisms, mollusk and coral biology and paleoclimatic issues have also been studied for a better understanding of the problematic of each case.



Bibliografía

1. P. M. Bellan. *Fundamentals of Plasma Physics*. Cambridge University Press, Cambridge, UK, 2008.
2. R. Fitzpatrick. *The Physics of Plasma*. Texas University Press, Texas, USA, 2008.
3. Schawlow, A. L. & Townes, C. H. Infrared and Optical Masers. *Phys. Rev.* **112**, 1940–1949 (1958).
4. Maiman, T. H. Stimulated Optical Radiation in Ruby. *Nature* **187**, 493–494 (1960).
5. Brech, F. y L. Cross. Optical microemission stimulated by a ruby maser. *Appl Spectrosc* **16**, 59 (1962).
6. J. A. Maxwell. The laser as a source in emission spectroscopy. *Chem Can* 10–11 (1963).
7. R. C. Rosan, F. Brech y D. Glick. Spectrographic analysis of nanogram samples by improved laser microprobe technique. *Fed Proc* **23**, 174 (1964).
8. E. F. Runge in: 10th Ottawa Symposium on Applied Spectroscopy (1963).
9. J. Debras-Guédon y N. Liodec. De l'utilisation du faisceau issu d'un amplificateur a ondes lumineuses par émission induite par rayonnement (laser à rubis), comme source énergétique pour l'excitation des spectres d'émission des éléments. *C R Hebd Seances Acad Sci* **257**, 3336–3339 (1963).
10. Loree, T. R. y Radziemski, L. J. Laser-induced breakdown spectroscopy: Time-integrated applications. *Plasma Chem. Plasma Process.* **1**, 271–279 (1981).
11. Hahn, D. W. y Omenetto, N. Laser-Induced Breakdown Spectroscopy (LIBS), Part I: Review of Basic Diagnostics and Plasma-Particle Interactions: Still-Challenging Issues Within the Analytical Plasma Community. *Appl. Spectrosc.* **64**, 335A–366A (2010).
12. Cremers, D. A. y Radziemski, L. J. *Handbook of laser-induced breakdown spectroscopy*. Jhon Wiley and sons, Chichester, UK, 2013.
13. Manzoor, S., Moncayo, S., Navarro-Villoslada, F., Ayala, J. A., Manuel de Villena, F. J y Caceres, J. O. Rapid identification and discrimination of bacterial



- strains by laser induced breakdown spectroscopy and neural networks. *Talanta* **121**, 65–70 (2014).
14. Anzano, J. M., Villoria, M. A., Ruíz-Medina, A. y Lasheras, R. J. Laser-induced breakdown spectroscopy for quantitative spectrochemical analysis of geological materials: Effects of the matrix and simultaneous determination. *Anal. Chim. Acta* **575**, 230–235 (2006).
 15. LI, H., LIU, M., CHEN, Z. y LI, R. Quantitative analysis of impurities in aluminum alloys by laser-induced breakdown spectroscopy without internal calibration. *Trans. Nonferrous Met. Soc. China* **18**, 222–226 (2008).
 16. Camacho, J. J., Díaz, L., Cid, J. P. y Poyato, J. M. L. Time-resolved spectroscopic diagnostic of the CO₂ plasma induced by a high-power CO₂ pulsed laser. *Spectrochim. Acta Part B At. Spectrosc.* **66**, 698–705 (2011).
 17. Piehler, T. N., Delucia, F. C., Munson, C. A. , Homan, B. E., Miziolek, A. W. y McNesby, K. L. Temporal evolution of the laser-induced breakdown spectroscopy spectrum of aluminum metal in different bath gases. *Appl. Opt.* **44**, 3654–3660 (2005).
 18. Aguilera, J. A., Aragón, C. y Peñalba, F. Plasma shielding effect in laser ablation of metallic samples and its influence on LIBS analysis. *Appl. Surf. Sci.* **127–129**, 309–314 (1998).
 19. De Giacomo, A., Dell’Aglia, M., Gaudiuso, R., Amoroso, S. y De Pascale, O. Effects of the background environment on formation, evolution and emission spectra of laser-induced plasmas. *Spectrochim. Acta Part B At. Spectrosc.* **78**, 1–19 (2012).
 20. Cremers, D. A. y Radziemski, L. J. *Handbook of Laser-Induced Breakdown Spectroscopy*. Jhon Wiley & Sons, Chichester, UK, 2006.
 21. Russo, R. E., Mao, X. L., Liu, C. y Gonzalez, J. Laser assisted plasma spectrochemistry: laser ablation. *J. Anal. At. Spectrom.* **19**, 1084–1089 (2004).
 22. Ageev, B., Ponomarev, Y. y Sapozhnikova, V. Laser Photoacoustic Detection of CO₂ in Old Disc Tree-Rings. *Sensors* **10**, 3305–3313 (2010).
 23. Alvira, F. C., Ramirez Rozzi, F. y Bilmes, G. M. Laser-Induced Breakdown Spectroscopy Microanalysis of Trace Elements in Homo sapiens Teeth. *Appl.*



- Spectrosc.* **64**, 313–319 (2010).
24. Bulgakova, N. M., Stoian, R., Rosenfeld, A., Hertel, I. V., Marine y W., Campbell, E. E. B.. A general continuum approach to describe fast electronic transport in pulsed laser irradiated materials: The problem of Coulomb explosion. *Appl. Phys. A* **81**, 345–356 (2005).
 25. Zhigilei, L. V., Lin, Z. y Ivanov, D. S. Atomistic Modeling of Short Pulse Laser Ablation of Metals: Connections between Melting, Spallation, and Phase Explosion†. *J. Phys. Chem. C* **113**, 11892–11906 (2009).
 26. Y. Raizer. *Laser-Induced Discharge Phenomena*. Springer, New York, USA, 1978.
 27. Capitelli, M., Casavola, A., Colonna, G. y De Giacomo, A. Laser-induced plasma expansion: theoretical and experimental aspects. *Spectrochim. Acta Part B At. Spectrosc.* **59**, 271–289 (2004).
 28. Colonna, G., Casavola, A. y Capitelli, M. Modelling of LIBS plasma expansion. *Spectrochim. Acta Part B At. Spectrosc.* **56**, 567–586 (2001).
 29. Elhassan, A., Giakoumaki, A., Anglos, D., Ingo, G. M., Robbiola, L. y Harith, M. A. Nanosecond and femtosecond Laser Induced Breakdown Spectroscopic analysis of bronze alloys. *Spectrochim. Acta Part B At. Spectrosc.* **63**, 504–511 (2008).
 30. Apitz, I. y Vogel, A. Material ejection in nanosecond Er:YAG laser ablation of water, liver, and skin. *Appl. Phys. A* **81**, 329–338 (2005).
 31. De Giacomo, A., Dell’Aglío, M., De Pascale, O. y Capitelli, M. From single pulse to double pulse ns-Laser Induced Breakdown Spectroscopy under water: Elemental analysis of aqueous solutions and submerged solid samples. *Spectrochim. Acta Part B At. Spectrosc.* **62**, 721–738 (2007).
 32. Koch, S., Garen, W., Müller, M. y Neu, W. Detection of chromium in liquids by laser induced breakdown spectroscopy (LIBS). *Appl. Phys. A* **79**, 1071–1073 (2004).
 33. Sarkar, A., Alamelu, D. y Aggarwal, S. K. Determination of thorium and uranium in solution by laser-induced breakdown spectrometry. *Appl. Opt.* **47**, G58–64 (2008).



34. Guirado, S., Fortes, F. J., Lazic, V. y Laserna, J. J. Chemical analysis of archeological materials in submarine environments using laser-induced breakdown spectroscopy. On-site trials in the Mediterranean Sea. *Spectrochim. Acta Part B At. Spectrosc.* **74–75**, 137–143 (2012).
35. Lazic, V., Colao, F., Fantoni, R. y Spizzicchino, V. Laser-induced breakdown spectroscopy in water: Improvement of the detection threshold by signal processing. *Spectrochim. Acta Part B At. Spectrosc.* **60**, 1002–1013 (2005).
36. Sakka, T., Masai, S., Fukami, K. y Ogata, Y. H. Spectral profile of atomic emission lines and effects of pulse duration on laser ablation in liquid. *Spectrochim. Acta Part B At. Spectrosc.* **64**, 981–985 (2009).
37. Wieser, P., Wurster, R. y Seiler, H. Identification of airborne particles by laser induced mass spectroscopy. *Atmospheric Environ.* 1967 **14**, 485–494 (1980).
38. Radziemski, L. J., Loree, T. R., Cremers, D. A. y Hoffman, N. M. Time-resolved laser-induced breakdown spectrometry of aerosols. *Anal. Chem.* **55**, 1246–1252 (1983).
39. Tjärnhage, T., Gradmark, P., Larsson, A., Mohamed, A., Landstrom, L., Sagerfors, E., Jonsson, P., Kullander, F. y Andersson, M. Development of a laser-induced breakdown spectroscopy instrument for detection and classification of single-particle aerosols in real-time. *Opt. Commun.* **296**, 106–108 (2013).
40. Dutouquet, C., Gallou, G., Le Bihan, O., Sirven, J. B., Dermigny, A., Torralba, B. y Frejafon, E. Monitoring of heavy metal particle emission in the exhaust duct of a foundry using LIBS. *Talanta* **127**, 75–81 (2014).
41. Bousquet, B., Sirven, J. B. y Canioni, L. Towards quantitative laser-induced breakdown spectroscopy analysis of soil samples. *Spectrochim. Acta Part B At. Spectrosc.* **62**, 1582–1589 (2007).
42. Bustamante, M. F., Rinaldi, C. A. y Ferrero, J. C. Laser induced breakdown spectroscopy characterization of Ca in a soil depth profile. *Spectrochim. Acta Part B At. Spectrosc.* **57**, 303–309 (2002).
43. Vadillo, J. M., Vadillo, I., Carrasco, F. y Laserna, J. J. Spatial distribution profiles of magnesium and strontium in speleothems using laser-induced



- breakdown spectrometry. *Fresenius J. Anal. Chem.* **361**, 119–123 (1998).
44. Marín-Roldán, A., Cruz, J. A., Martín-Chivelet, J., Turrero, M. J., Ortega, A. I., y Caceres, J. O. Evaluation of Laser Induced Breakdown Spectroscopy (LIBS) for detection of trace element variation through stalagmites potential for paleoclimate series reconstruction. *J. Appl. Laser Spectrosc.* **1**, 7–12 (2014).
 45. González, R., Lucena, P., Tobaría, L. M. y Laserna, J. J. Standoff LIBS detection of explosive residues behind a barrier. *J. Anal. At. Spectrom.* **24**, 1123–1126 (2009).
 46. Gottfried, J. L., Frank C. De Lucia, J., Munson, C. A. y Miziolek, A. W. Strategies for residue explosives detection using laser-induced breakdown spectroscopy. *J. Anal. At. Spectrom.* **23**, 205–216 (2008).
 47. Gottfried, J. L., Delucia, F., C. y Miziolek, A. W. Discrimination of explosive residues on organic and inorganic substrates using laser-induced breakdown spectroscopy. *J. Anal. At. Spectrom.* **24**, 288–296 (2009).
 48. Lucena, P., Gaona, I., Moros, J. y Laserna, J. J. Location and detection of explosive-contaminated human fingerprints on distant targets using standoff laser-induced breakdown spectroscopy. *Spectrochim. Acta Part B.* **85**, 71–77 (2013).
 49. Rehse, S. J. Laser-Based Identification of Pathogenic Bacteria. *Phys. Teach.* **47**, 152–156 (2009).
 50. Rehse, S. J., Mohaidat, Q. I. y Palchaudhuri, S. Towards the clinical application of laser-induced breakdown spectroscopy for rapid pathogen diagnosis: the effect of mixed cultures and sample dilution on bacterial identification. *Appl. Opt.* **49**, C27–C35 (2010).
 51. Rehse, S. J., Jeyasingham, N., Diedrich, J. y Palchaudhuri, S. A membrane basis for bacterial identification and discrimination using laser-induced breakdown spectroscopy. *J. Appl. Phys.* **105**, 102034–13 (2009).
 52. Multari, R. A., Cremers, D. A. y Bostian, M. L. Use of laser-induced breakdown spectroscopy for the differentiation of pathogens and viruses on substrates. *Appl. Opt.* **51**, B57–B64 (2012).
 53. Madhavi, M. Z., Wullschleger, S. D., Vass, A. A., Rodger, M. C. y Grissino-



- Mayer, H. High Resolution Laser-induced Breakdown Spectroscopy used in Homeland Security and Forensic Applications. *Spec. Issue First Indo-US Work. Spectrosc.* (2006).
54. Marín-Roldan, A., Manzoor, S., Moncayo, S., Navarro-Villoslada, F., Izquierdo, R. C. y Cáceres, J. O. Determination of the postmortem interval by Laser Induced Breakdown Spectroscopy using swine skeletal muscles. *Spectrochim. Acta Part B At. Spectrosc.* **88**, 186–191 (2013).
55. Moncayo, S., Manzoor, S., Ugidos, T., Navarro-Villoslada, F. y Cáceres, J. O. Discrimination of human bodies from bones and teeth remains by Laser Induced Breakdown Spectroscopy and Neural Networks. *Spectrochim. Acta Part B At. Spectrosc.* **101**, 21–25 (2014).
56. Tofanelli, M., Pardini, L., Borrini, M., Bartoli F., Bacci, A., Ulivo, A., Pitzalis, E., Mascherpa, M. C., Legnaioli, S., Lorenzetti, G., Pagnotta, S., Holanda Cavalcanti, G., Lezzerini, M. y Palleschi, V. Spectroscopic analysis of bones for forensic studies. *Spectrochim. Acta Part B At. Spectrosc.* **99**, 70–75 (2014).
57. Anzano, J., Lasheras, R.-J., Bonilla, B. y Casas, J. Classification of polymers by determining of C1:C2:CN:H:N:O ratios by laser-induced plasma spectroscopy (LIPS). *Polym. Test.* **27**, 705–710 (2008).
58. Anzano, J., Casanova, M.-E., Bermúdez, M.-S. Lasheras, R.-J. Rapid characterization of plastics using laser-induced plasma spectroscopy (LIPS). *Polym. Test.* **25**, 623–627 (2006).
59. Gaudiuso, R., Dell’Aglia, M., Pascale, O. D., Senesi, G. S. y Giacomo, A. D. Laser Induced Breakdown Spectroscopy for Elemental Analysis in Environmental, Cultural Heritage and Space Applications: A Review of Methods and Results. *Sensors* **10**, 7434–7468 (2010).
60. Harmon, R.S, DeLucia, F. C., McManus, E., McMillan, N. J., Jenkins, T. F., Walsh, M. E. y Miziolek, A. Laser-induced breakdown spectroscopy - An emerging chemical sensor technology for real-time field-portable, geochemical, mineralogical, and environmental applications. *Appl. Geochem.* **21**, 730–747 (2006).
61. Ortiz, P. Vázquez, M.A., Martín, J. M., Ctvrtnickova, T., Mateo, M. P. y



- Nicolas, G. Investigation of environmental pollution effects on stone monuments in the case of Santa Maria La Blanca, Seville (Spain). *Appl. Phys. A* **100**, 965–973 (2010).
62. Erdem, A. *et al.* Characterization of Iron age pottery from eastern Turkey by laser- induced breakdown spectroscopy (LIBS). *J. Archaeol. Sci.* **35**, 2486–2494 (2008).
63. Anglos, D., Couris, S. y Fotakis, C. Laser Diagnostics of Painted Artworks: Laser-Induced Breakdown Spectroscopy in Pigment Identification. *Appl. Spectrosc.* **51**, 1025–1030 (1997).
64. Pouli, P., Melessanaki, K., Giakoumaki, A., Argyropoulos, V. y Anglos, D. Measuring the thickness of protective coatings on historic metal objects using nanosecond and femtosecond laser induced breakdown spectroscopy depth profiling. *Spectrochim. Acta Part B*. **60**, 1163–1171 (2005).
65. Kumar, A., Yueh, F.-Y., Singh, J. P. y Burgess, S. Characterization of malignant tissue cells using laser induced breakdown spectroscopy in *Laser Applications to Chemical and Environmental Analysis* TuB₂ (Optical Society of America, 2004).
66. Samek, O. *et al.* Quantitative laser-induced breakdown spectroscopy analysis of calcified tissue samples. *Spectrochim. Acta Part B* **56**, 865–875 (2001).
67. Anzano, J. y Lasheras, R.-J. Strategies for the identification of urinary calculus by laser induced breakdown spectroscopy. *Talanta* **79**, 352–360 (2009).
68. US Department of Commerce, N. NIST Atomic Spectra Database. (2015). in <http://physics.nist.gov/PhysRefData/ASD/lines_form.html>
69. Marcos-Martinez, D., Ayala, J. A., Izquierdo-Hornillos, R. C., de Villena, F. J. M. y Caceres, J. O. Identification and discrimination of bacterial strains by laser induced breakdown spectroscopy and neural networks. *Talanta* **84**, 730–737 (2011).
70. Hussain, T. y Gondal, M. A. Monitoring and assessment of toxic metals in Gulf War oil spill contaminated soil using laser-induced breakdown spectroscopy. *Environ. Monit. Assess.* **136**, 391–399 (2007).
71. Lucilla Fornarini, R. F. Theoretical Modeling of Laser Ablation of Quaternary Bronze Alloys: Case Studies Comparing Femtosecond and Nanosecond LIBS



- Experimental Data. *J. Phys. Chem. A* **113**, 14364–74 (2009).
72. Thompson, J. R., Wiens, R. C., Barefield, J. E., Vaniman, D. T., Newsom, H. E. y Clegg, S. M. Remote laser-induced breakdown spectroscopy analyses of Dar al Gani 476 and Zagami Martian meteorites. *J. Geophys. Res. Planets* **111**, E05006 (2006).
73. Cremers, D. A., Yueh, F.-Y., Singh, J. P. y Zhang, H. in *Encyclopedia of Analytical Chemistry*, John Wiley & Sons, Chichester, UK, 2006.
74. Griem, H. R. *Principles of Plasma Spectroscopy*. (Cambridge University Press, 2005).
75. Bulajic, D., Corsi, M., Cristoforo, G., Legnaioli, S., Palleschi, V., Salvetti, A., y Tognoni, E. A procedure for correcting self-absorption in calibration free-laser induced breakdown spectroscopy. *Spectrochim. Acta Part B At. Spectrosc.* **57**, 339–353 (2002).
76. Moon, H.-Y., Herrera, K. K., Omenetto, N., Smith, B. W. y Winefordner, J. D. On the usefulness of a duplicating mirror to evaluate self-absorption effects in laser induced breakdown spectroscopy. *Spectrochim. Acta Part B At. Spectrosc.* **64**, 702–713 (2009).
77. De Giacomo, A., Dell’Aglío, M., De Pascale, O., Longo, S. & Capitelli, M. Laser induced breakdown spectroscopy on meteorites. *Spectrochim. Acta Part B At. Spectrosc.* **62**, 1606–1611 (2007).
78. Lasheras, R. J., Bello-Gálvez, C. y Anzano, J. M. Quantitative analysis of oxide materials by laser-induced breakdown spectroscopy with argon as an internal standard. *Spectrochim. Acta Part B At. Spectrosc.* **82**, 65–70 (2013).
79. Senesi, G. S., Dell’Aglío, M., De Giacomo, A., Zacccone, C., De Pascale, O., Miano, T. M. y Capitelli, M. Heavy metal concentrations in soils as determined by laser-induced breakdown spectroscopy (LIBS), with special emphasis on chromium. *Environ. Res.* **109**, 413–420 (2009).
80. A. Ciucci, M. C. New Procedure for Quantitative Elemental Analysis by Laser-Induced Plasma Spectroscopy. *Appl. Spectrosc. - Appl. Spectrosc.* **53**, 960–964 (1999).
81. Alvira, F. C., Orzi, D. J. O. y Bilmes, G. M. Surface treatment analyses of car



- bearings by using laser-induced breakdown spectroscopy. *Appl. Spectrosc.* **63**, 192–198 (2009).
82. Cardone, M. J., Palermo, P. J. y Sybrandt, L. B. Potential error in single-point-ratio calculations based on linear calibration curves with a significant intercept. *Anal. Chem.* **52**, 1187–1191 (1980).
83. Mao, X. L., Shannon, M. A., Fernandez, A. J. y Russo, R. E. Temperature and Emission Spatial Profiles of Laser-Induced Plasmas during Ablation Using Time-Integrated Emission Spectroscopy. *Appl. Spectrosc.* **49**, 1054–1062 (1995).
84. Simeonsson, J. B. y Miziolek, A. W. Time-resolved emission studies of ArF-laser-produced microplasmas. *Appl. Opt.* **32**, 939–947 (1993).
85. R.W.P. McWhirter. *Plasma Diagnostic Techniques*. (Academic Press, New York, 1995).
86. Cristoforetti, G., De Giacomo, A., Dell’Aglio, M., Legnaioli, S., Tognoni, E., Palleschi, V. y Omenetto, N. Local Thermodynamic Equilibrium in Laser-Induced Breakdown Spectroscopy: Beyond the McWhirter criterion. *Spectrochim. Acta Part B At. Spectrosc.* **65**, 86–95 (2010).
87. Zhao, X. Z., Shen, L. J., Lu, T. X. y Niemax, K. Spatial distributions of electron density in microplasmas produced by laser ablation of solids. *Appl. Phys. B* **55**, 327–330 (1992).
88. Grant, K. J. y Paul, G. L. Electron Temperature and Density Profiles of Excimer Laser-Induced Plasmas. *Appl. Spectrosc.* **44**, 1349–1354 (1990).
89. H.R. Griem. *Spectral Line Broadening by Plasmas*. (Academic Press, New York, 1974).
90. Lochte-Holtgreven, W. y Richter, J.. *Plasma diagnostics*. (North-Holland Pub. Co, 1968).
91. Aragón, C. y Aguilera, J. A. Characterization of laser induced plasmas by optical emission spectroscopy: A review of experiments and methods. *Spectrochim. Acta Part B At. Spectrosc.* **63**, 893–916 (2008).
92. Aguilera, J. A. y Aragón, C. Multi-element Saha–Boltzmann and Boltzmann plots in laser-induced plasmas. *Spectrochim. Acta Part B At. Spectrosc.* **62**, 378–



- 385 (2007).
93. Camacho, J. J., Díaz, L., Santos, M. y Poyato, J. M. L. Time-resolved optical emission spectroscopic measurements of He plasma induced by a high-power CO₂ pulsed laser. *Spectrochim. Acta Part B At. Spectrosc.* **66**, 57–66 (2011).
 94. Hornáček, M., Hornáček, M., Rakovský, J., Hudec, P. y Veis, P. Determination of Si/Al molar ratios in microporous zeolites using calibration-free laser induced breakdown spectroscopy. *Spectrochim. Acta Part B At. Spectrosc.* **88**, 69–74 (2013).
 95. Moncayo, S., Manzoor, S., Navarro-Villoslada, F. y Cáceres, J. O. Evaluation of supervised chemometric methods for sample classification by Laser Induced Breakdown Spectroscopy. *Chemom. Intell. Lab. Syst.* **146**, 354–364 (2015).
 96. Currie, L. A. Nomenclature in evaluation of analytical methods including detection and quantification capabilities (IUPAC Recommendations 1995). *Pure Appl. Chem.* **67**, (1995).
 97. Van Der Voet, H. Encyclopedia of Environmetrics Detection Limits. Encyclopedia of Environmetrics. 1. (John Wiley & Sons, Ltd, 2006).
 98. Currie, L. A. Detection and quantification limits: origins and historical overview. *Anal. Chim. Acta* **391**, 127–134 (1999).
 99. Vessman, J. *et al.* Selectivity in analytical chemistry (IUPAC Recommendations 2001). *Pure Appl. Chem.* **73**, (2001).
 100. Brown, S. D. in *Encyclopedia of Analytical Chemistry* John Wiley & Sons, Chichester, UK, 2006.
 101. Gemperline, P. *Practical Guide To Chemometrics*. CRC Press, Florida, USA, 2006.
 102. Brown, S. D., Tauler, R. y Walczak, B. *Comprehensive Chemometrics: Chemical and Biochemical Data Analysis*. (Elsevier, Amsterdam, 2009).
 103. Kuncheva, L. I. *Combining Pattern Classifiers: Methods and Algorithms*. Wiley-Interscience, Hoboken, NJ, USA, 2004.
 104. Ballabio, D. y Todeschini, R. in *Infrared Spectroscopy for Food Quality Analysis and Control* 83–104 Academic Press, New York, USA, 2009.
 105. K. Pearson. On lines and planes of closest fit to systems of points in space.



- Philos. Mag.* **2**, 559–572 (1901).
106. Hotelling, H. Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* **24**, 417–441 (1933).
107. Brereton, R. G. *Chemometrics for Pattern Recognition*. (John Wiley & Sons Ltd., Hoboken, NJ, 2009).
108. Bro, R. & Smilde, A. K. Principal component analysis. *Anal. Methods* **6**, 2812–2831 (2014).
109. Max Welling. *Fisher Linear Discriminant Analysis* University of Toronto, Toronto, Toronto, Canada, 2008.
110. De Maesschalck, R., Jouan-Rimbaud, D. y Massart, D. L. The Mahalanobis distance. *Chemom. Intell. Lab. Syst.* **50**, 1–18 (2000).
111. Kim, H.-C., Kim, D. y Bang, S. Y. Extensions of LDA by PCA mixture model and class-wise features. *Pattern Recognit.* **36**, 1095–1105 (2003).
112. Wold, S. y Sjöström, M. SIMCA: A Method for Analyzing Chemical Data in Terms of Similarity and Analogy. *Chemom. Theory Appl.* **52**, 243–282 (1977).
113. Myakalwar, A. K., Sreedhar, S., Barman, I., Dingari, N. C., Kiran, P. P. y Manoj, K. G. Laser-induced breakdown spectroscopy-based investigation and classification of pharmaceutical tablets using multivariate chemometric analysis. *Talanta* **87**, 53–59 (2011).
114. Pontes, M. J. C., Cortez, J., Galvao, R. K., Coelho, R. MM., Chiba, M. K., de Abreu, M. F. y Madai, B. E. Classification of Brazilian soils by using LIBS and variable selection in the wavelet domain. *Anal. Chim. Acta* **642**, 12–18 (2009).
115. Sirven, J. B., Salle, B., Mauchien, P., Lacour, J., Maurice, S. y Manhes, G. Feasibility study of rock identification at the surface of Mars by remote laser-induced breakdown spectroscopy and three chemometric methods. *J. Anal. At. Spectrom.* **22**, 1471–1480 (2007).
116. Brereton, R. G. y Lloyd, G. R. Partial least squares discriminant analysis: taking the magic away. *J. Chemom.* **28**, 213–225 (2014).
117. Barker, M. y Rayens, W. Partial least squares for discrimination. *J. Chemom.* **17**, 166–173 (2003).
118. Pérez, N. F., Ferré, J. y Boqué, R. Calculation of the reliability of classification



- in discriminant partial least-squares binary classification. *Chemom. Intell. Lab. Syst.* **95**, 122–128 (2009).
119. Senesi, G. S. Laser-Induced Breakdown Spectroscopy (LIBS) applied to terrestrial and extraterrestrial analogue geomaterials with emphasis to minerals and rocks. *Earth-Sci. Rev.* **139**, 231–267 (2014).
120. Tian, Y., Wang, Z., Han, X., Hou, H. y Zheng, R. Comparative investigation of partial least squares discriminant analysis and support vector machines for geological cuttings identification using laser-induced breakdown spectroscopy. *Spectrochim. Acta Part B* **102**, 52–57 (2014).
121. Martin, M. Z., Mayes, M. A., Heal, K. R., Brice, D. J. y Wulfschleger, S. D. Investigation of laser-induced breakdown spectroscopy and multivariate analysis for differentiating inorganic and organic C in a variety of soils. *Spectrochim. Acta Part B* **87**, 100–107 (2013).
122. L. Breiman, J. Friedman, C.J. Stone y R.A. Olshen. *Classification and regression trees*. Wadsworth Inc, Ohio, USA, 1984.
123. Rousseau, R., Govaerts, B., Verleysen, M. y Boulanger, B. Comparison of some chemometric tools for metabonomics biomarker identification. *Chemom. Intell. Lab. Syst.* **91**, 54–66 (2008).
124. R.L. Lawrence, A. W. Rule-Based Classification Systems Using Classification and Regression Tree (CART) Analysis. *Photogramm. Eng. Remote Sens.* **67**, (2001).
125. McCullagh, P. y Nelder, J. A. *Generalized Linear Models, Second Edition*. (Chapman and Hall, London, 1989).
126. Boser, B. E., Guyon, I. M. y Vapnik, V. N. A Training Algorithm for Optimal Margin Classifiers. in *Proceedings of the Fifth Annual Workshop on Computational Learning Theory* 144–152 ACM, USA, 1992.
127. Scholkopf, B. & Smola, A. J. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. The MIT Press, Masachuset, USA 2001.
128. Vapnik, V. N. *The Nature of Statistical Learning Theory*. Springer, New York, USA, 2000.



129. Cortes, C. y Vapnik, V. Support-Vector Networks. *Mach. Learn.* **20**, 273–297 (1995).
130. Bouchaffra, D., Cheriet, M., Jodoin, P.-M. y Beck, D. Machine learning and pattern recognition models in change detection. *Pattern Recognit.* **48**, 613–615 (2015).
131. Wang, G., Lam, K.-M., Deng, Z. y Choi, K.-S. Prediction of mortality after radical cystectomy for bladder cancer by machine learning techniques. *Comput. Biol. Med.* **63**, 124–132 (2015).
132. Vogel, A. C., Church, J. A., Power, J. D., Miezin, F. M., Petersr, S. E. y Schlaggar, G. L. Functional network architecture of reading-related regions across development. *Brain Lang.* **125**, 231–243 (2013).
133. Sporns, O. Cerebral cartography and connectomics. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **370**, 20140173 (2015).
134. Bishop, C. M. *Neural Networks for Pattern Recognition*. Oxford University Press, Oxford, USA, 1996.
135. Demuth, H. B., Beale, M. H. y Hagan, M. T. *Neural network toolbox for use with MATLAB: User's guide 9th for version 6.0 (Release 2008a)*. Math Works, Natick, USA, 2007.
136. Gelenbe, E. Random Neural Networks with Negative and Positive Signals and Product Form Solution. *Neural Comput.* **1**, 502–510 (1989).
137. Gelenbe, E. Learning in the Recurrent Random Neural Network. *Neural Comput.* **5**, 154–164 (1993).
138. Smolensky, P. Information Processing in Dynamical Systems: Foundations of Harmony Theory ; CU-CS-321-86. *Comput. Sci. Tech. Rep.* (1986).
139. Hopfield, J. J. & Tank, D. W. Computing with neural circuits: A model. *Science* **233**, 625–633 (1986).
140. Berglund, M., Raiko, T. & Cho, K. Measuring the usefulness of hidden units in Boltzmann machines with mutual information. *Neural Netw.* **64**, 12–18 (2015).
141. Block, H. D. A review of “perceptrons: An introduction to computational geometry *Inf. Control* **17**, 501–522 (1970).
142. Ding, S., Li, H., Su, C., Yu, J. y Jin, F. Evolutionary artificial neural networks: a



- review. *Artif. Intell. Rev.* **39**, 251–260 (2011).
143. Falco, I. D., Cioppa, A. D., Natale, P. y Tarantino, E. in *Soft Computing in Engineering Design and Manufacturing* 3–12 Springer, London, UK, 1998.
144. Rumelhart, D. E., Hinton, G. E. y Williams, R. J. Learning representations by back-propagating errors. *Nature* **323**, 533–536 (1986).
145. Fritsch, J., Dr, P., Waibel, A. y Finke, M. *Modular Neural Networks for Speech Recognition*. Carnellie Mellon University, Pittsburgh, USA, 1996.
146. Sanger, T. D. Optimal unsupervised learning in a single-layer linear feedforward neural network. *Neural Netw.* **2**, 459–473 (1989).
147. Moncayo, S., Manzoor, S., Ugidos, T., Navarro-Villoslada, F. y Caceres, J. O. Discrimination of human bodies from bones and teeth remains by Laser Induced Breakdown Spectroscopy and Neural Networks. *Spectrochim. Acta Part B At. Spectrosc.* **101**, 21–25 (2014).
148. Wilson, E. Backpropagation Learning for Systems with Discrete-Valued Functions. *Proceedings of the World Congress on Neural Networks*. Stanford University, Standfor, USA, 1994.
149. Liu, Z., Liu, A., Wang, C. y Niu, Z. Evolving Neural Network Using Real Coded Genetic Algorithm (GA) for Multispectral Image Classification. *Future Gener Comput Syst* **20**, 1119–1129 (2004).
150. Kolen, J. F. y Pollack, J. B. Back Propagation is Sensitive to Initial Conditions. in *Proceedings of the 1990 Conference on Advances in Neural Information Processing Systems 3* 860–867. Morgan Kaufmann Publishers Inc, USA, 1990.
151. Laser Facts. Prens Article
<http://www.nobelprize.org/educational/physics/laser/facts/applications.html>
152. Coherent, Inc. Lasers: Understanding the Basic. *Photonics Handbook*
153. G. Overton y A. Noguee. Laser Marketplace 2014: Lasers forge 21st century innovations. (2014). at <http://www.laserfocusworld.com/articles/print/volume-50/issue-01/features/laser-marketplace-2014-lasers-forge-21st-century-innovations.html>
154. Townes, C. H., Básov, N. y Prójorov, P. Nobel Prize in Physics 1964 - Presentation Speech for Nobel Prize. (1964).



155. Camacho, J. J., Díaz, L., Martínez-Ramírez, S. y Cáceres, J. O. Time-and space-resolved spectroscopic characterization of laser-induced swine muscle tissue plasma. *Spectrochim. Acta - Part B At. Spectrosc.* **111**, 92–101 (2015).
156. Camacho, J. J., Díaz, L., Marín-Roldán, A., Moncayo, S y Cáceres, J. O. Plume dynamics of laser-produced swine muscle tissue plasma. *Appl. Spectrosc.* (2016).
157. Telle, H. H., Ureña, A. G. y Donovan, R. J. *Laser Chemistry: Spectroscopy, Dynamics and Applications*. Wiley & Sons, Chichester, UK, 2007.
158. Arai, A., Harada, Y. y Tateno, R. Czerny-turner spectroscope. (2003) at <http://www.google.com.ar/patents/US6507398>
159. R. L. Hiliard. Imaging spectrograph for multiorder spectroscopy at <http://www.google.com/patents/US6628383>
160. Hornáčeková, M., Gromusová, Z., M., Hornáček, M., Rakovský, J. Hudec, P. y Veis, P. Calibration analysis of zeolites by laser induced breakdown spectroscopy. *Spectrochim. Acta Part B At. Spectrosc.* **74–75**, 119–123 (2012).
161. Hornáčeková, M., Hornáček, M., Rakovský, J., Hudec, P. y Veis, P. Determination of Si/Al molar ratios in microporous zeolites using calibration-free laser induced breakdown spectroscopy. *Spectrochim. Acta Part B At. Spectrosc.* **88**, 69–74 (2013).
162. Rodríguez Carnero, P., Calvo Sánchez, M., Molina Collado, Z., Durán Moreno, J. y Martínez Sanz, N. El síndrome del aceite tóxico: 30 años después. *Rev. Esp. Med. Leg.* **37**, 155–161 (2011).
163. Beauchamp, G. K., Russell, S. J., Keast, Morel, D., Lin, J. Pika, J., Han Q., Ho Lee, C., Amos, B., Paul, A. S. Breslin Phytochemistry: Ibuprofen-like activity in extra-virgin olive oil. *Nature* **437**, 45–46 (2005).
164. Gimeno, E., Castellote, A. I., Lamuela-Raventós, R. M., De la Torre, M. C. y López-Sabater, M. C. The effects of harvest and extraction methods on the antioxidant content (phenolics, α -tocopherol, and β -carotene) in virgin olive oil. *Food Chem.* **78**, 207–211 (2002).
165. Benincasa, C., Lewis, J., Perri, E., Sindona, G. y Tagarelli, A. Determination of trace element in Italian virgin olive oils and their characterization according to



- geographical origin by statistical analysis. *Anal. Chim. Acta* **585**, 366–370 (2007).
166. Peña, F., Cárdenas, S., Gallego, M. y Valcárcel, M. Direct olive oil authentication: Detection of adulteration of olive oil with hazelnut oil by direct coupling of headspace and mass spectrometry, and multivariate regression techniques. *J. Chromatogr. A* **1074**, 215–221 (2005).
167. Papadopoulos, K., Triantis, T., Tzikis, C. H., Nikokavoura, A. y Dimotikali, D. Investigations of the adulteration of extra virgin olive oils with seed oils using their weak chemiluminescence. *Anal. Chim. Acta* **464**, 135–140 (2002).
168. Hajimahmoodi, M., Vander Heyden, Y., Sadeghi, N., Jannat, B., Oveisi, M. R., Shahbazian, S. Gas-chromatographic fatty-acid fingerprints and partial least squares modeling as a basis for the simultaneous determination of edible oil mixtures. *Talanta* **66**, 1108–1116 (2005).
169. Christopoulou, E., Lazaraki, M., Komaitis, M. y Kaselimis, K. Effectiveness of determinations of fatty acids and triglycerides for the detection of adulteration of olive oils with vegetable oils. *Food Chem.* **84**, 463–474 (2004).
170. Zabarar, D. y Gordon, M. H. Detection of pressed hazelnut oil in virgin olive oil by analysis of polar components: improvement and validation of the method. *Food Chem.* **84**, 475–483 (2004).
171. Saba, A., Mazzini, F., Raffaelli, A., Mattei, A. y Salvadori, P. Identification of 9(E),11(E)-18:2 Fatty Acid Methyl Ester at Trace Level in Thermal Stressed Olive Oils by GC Coupled to Acetonitrile CI-MS and CI-MS/MS, a Possible Marker for Adulteration by Addition of Deodorized Olive Oil. *J. Agric. Food Chem.* **53**, 4867–4872 (2005).
172. Vigli, G., Philippidis, A., Spyros, A. y Dais, P. Classification of Edible Oils by Employing ³¹P and ¹H NMR Spectroscopy in Combination with Multivariate Statistical Analysis. A Proposal for the Detection of Seed Oil Adulteration in Virgin Olive Oils. *J. Agric. Food Chem.* **51**, 5715–5722 (2003).
173. Tay, A., Singh, R. K., Krishnan, S. S. y Gore, J. P. Authentication of Olive Oil Adulterated with Vegetable Oils Using Fourier Transform Infrared Spectroscopy. *Lebensm.-Wiss. -Technol.* **35**, 99–103 (2002).



174. Ozen, B. F. y Mauer, L. J. Detection of Hazelnut Oil Adulteration Using FT-IR Spectroscopy. *J. Agric. Food Chem.* **50**, 3898–3901 (2002).
175. Gurdeniz, G. y Ozen, B. Detection of adulteration of extra-virgin olive oil by chemometric analysis of mid-infrared spectral data. *Food Chem.* **116**, 519–525 (2009).
176. Baeten, V., Meurens, M., Morales, M. T. y Aparicio, R. Detection of Virgin Olive Oil Adulteration by Fourier Transform Raman Spectroscopy. *J. Agric. Food Chem.* **44**, 2225–2230 (1996).
177. Zou, M.Q., Xiao-Fang, Z., Xiao-Hua, Q., Han-Lun, M., Ying, D., Chun-Wei L., Xun, G. y Hong W. Rapid Authentication of Olive Oil Adulteration by Raman Spectrometry. *J. Agric. Food Chem.* **57**, 6001–6006 (2009).
178. Bartlett, J. G. y Mahon, J. H. Identification of oils and the detection of oil adulteration by differential IR spectroscopy. *J Assoc Anal Chem* **41**, 450–459 (1958).
179. Li-Chan, E. Developments in the detection of adulteration of olive oil. *Trends Food Sci. Technol.* **5**, 3–11 (1994).
180. Lopez-Diez, E. C., Bianchi, G. y Goodacre, R. Rapid Quantitative Assessment of the Adulteration of Virgin Olive Oils with Hazelnut Oils Using Raman Spectroscopy and Chemometrics. *J. Agric. Food Chem.* **51**, 6145–6150 (2003).
181. Korifi, R., Le Dréau, Y., Molinet, J., Artaud, J. y Dupuy, N. Composition and authentication of virgin olive oil from French PDO regions by chemometric treatment of Raman spectra. *J. Raman Spectrosc.* **42**, 1540–1547 (2011).
182. Dong, W., Zhang, Y., Zhang, B. y Wang, X. Quantitative analysis of adulteration of extra virgin olive oil using Raman spectroscopy improved by Bayesian framework least squares support vector machines. *Anal. Methods* **4**, 2772–2777 (2012).
183. Beattie, J. R., Bell, S. E. J. y Moss, B. W. A critical evaluation of Raman spectroscopy for the analysis of lipids: fatty acid methyl esters. *Lipids* **39**, 407–419 (2004).
184. Breitman, M., Ruiz-Moreno, S. y Gil, A. L. Experimental problems in Raman spectroscopy applied to pigment identification in mixtures. *Spectrochim. Acta.*



- A. Mol. Biomol. Spectrosc.* **68**, 1114–1119 (2007).
185. Hahn, D. W. y Omenetto, N. Laser-Induced Breakdown Spectroscopy (LIBS), Part II: Review of Instrumental and Methodological Approaches to Material Analysis and Applications to Different Fields. *Appl. Spectrosc.* **66**, 347–419 (2012).
186. Miziolek, A. W., Palleschi, V. y Schechter, I. *Laser Induced Breakdown Spectroscopy (LIBS)*. Cambridge University Press, Cambridge, UK, 2008.
187. Caceres, J. O., Tornero López, J., Telle, H. H. y González Ureña, A. Quantitative analysis of trace metal ions in ice using laser-induced breakdown spectroscopy. *Spectrochim. Acta Part B At. Spectrosc.* **56**, 831–838 (2001).
188. Lasheras, R. J., Bello-Gálvez, C. y Anzano, J. Identification of polymers by libs using methods of correlation and normalized coordinates. *Polym. Test.* **29**, 1057–1064 (2010).
189. Tripathi, M. M., Srinivasan, K. K., Krishnan, S. R., Yueh, F.-Y. y Singh, J. P. A comparison of multivariate LIBS and chemiluminescence-based local equivalence ratio measurements in premixed atmospheric methane–air flames. *Fuel*. **93**, 684–691 (2012).
190. Ramil, A., López, A. J. y Yáñez, A. Application of artificial neural networks for the rapid classification of archaeological ceramics by means of laser induced breakdown spectroscopy (LIBS). *Appl. Phys. A* **92**, 197–202 (2008).
191. Sirven, J.B., Bousquet, B., Canioni, L., Sarger, L., Tellier, S., Potin-Gautie, M. y Le Hecho, I. Qualitative and quantitative investigation of chromium-polluted soils by laser-induced breakdown spectroscopy combined with neural networks analysis. *Anal. Bioanal. Chem.* **385**, 256–262 (2006).
192. Møller, M. F. A scaled conjugate gradient algorithm for fast supervised learning. *Neural Netw.* **6**, 525–533 (1993).
193. Bishop, C. M. *Pattern Recognition and Machine Learning*. Springer, New York, USA, 2007.
194. Zhang, J., Ma, G., Zhu, H., Xi, J. y Ji, Z. Accurate quantitative analysis of metal oxides by laser-induced breakdown spectroscopy with a fixed plasma temperature calibration method. *J. Anal. At. Spectrom.* **27**, 1903–1908 (2012).



195. Frydenvang, J., Kinch, K. M., Husted, S. y Madsen, M. B. An Optimized Calibration Procedure for Determining Elemental Ratios Using Laser-Induced Breakdown Spectroscopy. *Anal. Chem.* **85**, 1492–1500 (2013).
196. Sarzyski, A., Skrzeczanowski, W. Marczak, J. Physical phenomena disturbing LIBS analysis. in *Proceedings of SPIE* Military University of Technology, Poland, 2007.
197. Nordstrom, R. J. Study of Laser-Induced Plasma Emission Spectra of N₂, O₂, and Ambient Air in the Region 350 nm to 950 nm. *Appl. Spectrosc.* **49**, 1490–1499 (1995).
198. Aceto, M., Abollino, O., Bruzzoniti, M. C., Mentasti, E., Sarzanini, C. y Malandrino, M. Determination of metals in wine with atomic spectroscopy (flame-AAS, GF-AAS and ICP-AES); a review. *Food Addit. Contam.* **19**, 126–133 (2002).
199. Larsen, F. H., van den Berg, F. y Engelsens, S. B. An exploratory chemometric study of ¹H NMR spectra of table wines. *J. Chemom.* **20**, 198–208 (2006).
200. Brescia, M., Caldarola, V., De Giglio, A., Benedetti, D., Fanizzi, F. P. y Sacco, A. Characterization of the geographical origin of Italian red wines based on traditional and nuclear magnetic resonance spectrometric determinations. *Anal. Chim. Acta* **458**, 177–186 (2002).
201. Makris, D. P., Kallithraka, S. y Mamalos, A. Differentiation of young red wines based on cultivar and geographical origin with application of chemometrics of principal polyphenolic constituents. *Talanta* **70**, 1143–1152 (2006).
202. Caceres, J. O., Moncayo, S., Rosales, J. d., Manuel de villena, F. J., Alvira, F. C. y Bilmes, G. M. Application of Laser-Induced Breakdown Spectroscopy (LIBS) and Neural Networks to Olive Oils Analysis. *Appl. Spectrosc.* **67**, 1064–1072 (2013).
203. Sirven, J. B., Bousquet, B., Canioni, L. y Sarger, L. Laser-Induced Breakdown Spectroscopy of Composite Samples: Comparison of Advanced Chemometrics Methods. *Anal. Chem.* **78**, 1462–1469 (2006).
204. Cámara, M., Torrecilla, J. S., Caceres, J. O., Sánchez Mata, M. C. y Fernández-Ruiz, V. Neural Network Analysis of Spectroscopic Data of Lycopene and β -



- Carotene Content in Food Samples Compared to HPLC-UV-Vis. *J. Agric. Food Chem.* **58**, 72–75 (2010).
205. T. Dale Stewart. *Essentials of Forensic Anthropology, Especially As Developed in the United States*. Charles C. Thomas Pub Ltd, Illinois, USA, 1979.
206. Interpol. INTERPOL Disaster Victim Identification Guide. (2014).
207. Beauthier, J. P., Edd y De Valck, Lefevre, P. y De Winne, J. Mass Disaster Victim Identification: The Tsunami Experience. *Open Forensic Sci. J.* **2**, 54–62 (2009).
208. Sweet, D. Solving certain dental records problems with technology--the Canadian solution in the Thailand tsunami response. *Forensic Sci. Int.* **159**, S20–23 (2006).
209. Blau, S. y Briggs, C. A. The role of forensic anthropology in Disaster Victim Identification (DVI). *Forensic Sci. Int.* **205**, 29–35 (2011).
210. Hill, A. J., Hewson, I. y Lain, R. The role of the forensic odontologist in disaster victim identification: lessons for management. *Forensic Sci. Int.* **205**, 44–47 (2011).
211. Anđelinović, Š. Ivkošić, A., Pačić, F., Rezić, B., Definis-Gojanovic, M. y Primorac, D. Twelve-year experience in identification of skeletal remains from mass graves. *Croat Med J* **46**, 530–539 (2005).
212. Gunby, P. Medical Team Seeks to Identify Human Remains From Mass Graves of War in Former Yugoslavia. *JAMA J. Am. Med. Assoc.* **272**, 1804 –1806 (1994).
213. Castro, W., Hoogewerff, J., Latkoczy, C. y Almirall, J. R. Application of laser ablation (LA-ICP-SF-MS) for the elemental analysis of bone and teeth samples for discrimination purposes. *Forensic Sci. Int.* **195**, 17–27 (2010).
214. Fortes, F. J., Moros, J., Lucena, P., Cabalin, L. M. y Laserna, J. J. Laser-Induced Breakdown Spectroscopy. *Anal. Chem.* **85** (2), 640–669 (2013).
215. Myakalwar, A. K. *et al.* Laser-induced breakdown spectroscopy-based investigation and classification of pharmaceutical tablets using multivariate chemometric analysis. *Talanta* **87**, 53–59 (2011).
216. Turunen, M. J., Prantner, V., Jurvelin, J. S., Kröger, H. y Isaksson, H.



- Composition and microarchitecture of human trabecular bone change with age and differ between anatomical locations. *Bone* **54**, 118–125 (2013).
217. Higgins, D. y Austin, J. J. Teeth as a source of DNA for forensic identification of human remains: A Review. *Sci. Justice* **53**, 433–441 (2013).
218. Mundorff, A. Z., Bartelink, E. J. y Mar-Cash, E. DNA Preservation in Skeletal Elements from the World Trade Center Disaster: Recommendations for Mass Fatality Management*. *J. Forensic Sci.* **54**, 739–745 (2009).
219. Mundorff, A. Z., Shaler, R., Bieschke, E. y Mar-Cash, E. in *Recovery, Analysis, and Identification of Commingled Human Remains* 285–299 Humana Press, New York, USA, 2008.
220. Budimlija, Z. M. *et al.* World Trade Center human identification project: experiences with individual body identification cases. *Croat. Med. J.* **44**, 259–263 (2003).
221. Brereton, R. G. *Applied Chemometrics for Scientists*. (John Wiley & Sons Ltd., New York, 2007).
222. Ballabio, D. y Todeschini, R. in *Infrared Spectroscopy for Food Quality Analysis and Control*. Chapter 4: 83–104. Elsevier Academic Press, Oxford, UK., 2009.
223. Yueh, F.-Y., Zheng, H., Singh, J. P. y Burgess, S. Preliminary evaluation of laser-induced breakdown spectroscopy for tissue classification. *Spectrochim. Acta Part B At. Spectrosc.* **64**, 1059–1067 (2009).
224. Vítková, G. *et al.* Fast identification of biominerals by means of stand-off laser-induced breakdown spectroscopy using linear discriminant analysis and artificial neural networks. *Spectrochim. Acta Part B At. Spectrosc.* **73**, 1–6 (2012).
225. Dingari, N. C., Barman, I., Myakalwar, A. K., Tewari, S. P. y Kumar Gundawar, M. Incorporation of Support Vector Machines in the LIBS Toolbox for Sensitive and Robust Classification Amidst Unexpected Sample and System Variability. *Anal. Chem.* **84**, 2686–2694 (2012).
226. Gottfried, J. L. Discrimination of biological and chemical threat simulants in residue mixtures on multiple substrates. *Anal. Bioanal. Chem.* **400**, 3289–3301 (2011).



227. Ballabio, D. y Consonni, V. Classification tools in chemistry. Part 1: linear models. PLS-DA. *Anal. Methods* **5**, 3790–3798 (2013).
228. Jiang, Y., Cukic, B. y Ma, Y. Techniques for evaluating fault prediction models. *Empir. Softw. Eng.* **13**, 561–595 (2008).
229. Cordella, C. B. Y. y Bertrand, D. SAISIR: A new general chemometric toolbox. *TrAC Trends Anal. Chem.* **54**, 75–82 (2014).
230. Chen, T. *et al.* Increasing temperature anomalies reduce coral growth in the Weizhou Island, northern South China Sea. *Estuar. Coast. Shelf Sci.* **130**, 121–126 (2013).
231. J. Ring, M., Lindner, D., F. Cross, E. y E. Schlesinger, M. Causes of the Global Warming Observed since the 19th Century. *Atmospheric Clim. Sci.* **02**, 401–415 (2012).
232. Gutierrez Zugasti, I., Garcia Escarzaga, A., Martin-Chivelet, J. y Gonzalez-Morales, M. R. Determination of sea surface temperatures using oxygen isotope ratios from *Phorcus lineatus* (Da Costa 1778), in northern Spain: implications for paleoclimate and archaeological studies. The Holocene. *The Holocene.* **24**, 1002–1014 (2015).
233. Burton, I., Challenger, B., Huq, S., Klein, R. J. T. y Yohe, G. *Adaptation to Climate Change in the Context of Sustainable Development and Equity.*
234. Álvarez, M., Briz Godino, I., Balbo, A. y Madella, M. Shell middens as archives of past environments, human dispersal and specialized resource management. *Quat. Int.* **239**, 1–7 (2011).
235. Andrus, C. F. T. Shell midden sclerochronology. *Quat. Sci. Rev.* **30**, 2892–2905 (2011).
236. Colonese, A. C., Troelstra, S., Ziveri, P., Lo Vetro, D. y Tommasini, S. Mesolithic shellfish exploitation in SW Italy: seasonal evidence from the oxygen isotopic composition of *Osilinus turbinatus* shells. *J. Archaeol. Sci.* **36**, 1935–1944 (2009).
237. Prendergast, A. L., Azzopardi, M., O'Connell, T. C., Hunt, C., Barker, G. y Stecens, R. E. Oxygen isotopes from *Phorcus* (*Osilinus*) *turbinatus* shells as a proxy for sea surface temperature in the central Mediterranean: A case study



- from Malta. *Chem. Geol.* **345**, 77–86 (2013).
238. Wang, T., Surge, D. y Mithen, S. Seasonal temperature variability of the Neoglacial (3300–2500 BP) and Roman Warm Period (2500–1600 BP) reconstructed from oxygen isotope ratios of limpet shells (*Patella vulgata*), Northwest Scotland. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **317–318**, 104–113 (2012).
239. Mannino, M. A., Thomas, K. D., Leng, M. J. y Sloane, H. J. Shell growth and oxygen isotopes in the topshell *Osilinus turbinatus*: resolving past inshore sea surface temperatures. *Geo-Mar. Lett.* **28**, 309–325 (2008).
240. Ferguson, J. E., Henderson, G. M., Fa, D. A., Finlayson, J. C. y Charnley, N. R. Increased seasonality in the Western Mediterranean during the last glacial from limpet shell geochemistry. *Earth Planet. Sci. Lett.* **308**, 325–333 (2011).
241. Freitas, P. S., Clarke, L. J., Kennedy, H. y Richardson, C. A. The potential of combined Mg/Ca and $\delta^{18}\text{O}$ measurements within the shell of the bivalve *Pecten maximus* to estimate seawater $\delta^{18}\text{O}$ composition. *Chem. Geol.* **291**, 286–293 (2012).
242. Elliot, M. Welsh, K., Chilcott, C., McCulloch, M., Chappell, J. y Ayling, B. Profiles of trace elements and stable isotopes derived from giant long-lived *Tridacna gigas* bivalves: Potential applications in paleoclimate studies. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **280**, 132–142 (2009).
243. Lazareth, C. E., Guzman, N., Poitrasson, F., Candaudap, F. y Ortlieb, L. Nyctemeral variations of magnesium intake in the calcitic layer of a Chilean mollusk shell (*Concholepas concholepas*, Gastropoda). *Geochim. Cosmochim. Acta* **71**, 5369–5383 (2007).
244. Takesue, R. K. y van Geen, A. Mg/Ca, Sr/Ca, and stable isotopes in modern and Holocene *Protothaca staminea* shells from a northern California coastal upwelling region. *Geochim. Cosmochim. Acta* **68**, 3845–3861 (2004).
245. Surge, D. y Lohmann, K. C. Evaluating Mg/Ca ratios as a temperature proxy in the estuarine oyster, *Crassostrea virginica*. *J. Geophys. Res. Biogeosciences* **113**, G02001 (2008).
246. Carré, M., Bentaleb, O., Bruguier, O., Ordinola, E., Barrett, N. y Fontugne, M.



- Calcification rate influence on trace element concentrations in aragonitic bivalve shells: Evidences and mechanisms. *Geochim. Cosmochim. Acta* **70**, 4906–4920 (2006).
247. Ford, H. L., Schllenberg, S. A., Becker, B. J., Deutschman, D. L., Dyck, K. A. y Koch, P. L. Evaluating the skeletal chemistry of *Mytilus californianus* as a temperature proxy: Effects of microenvironment and ontogeny. *Paleoceanography* **25**, PA1203 (2010).
248. Foster, L. C., Finch, A. A., Allison, N., Andersson, C. y Clarke, L. J. Mg in aragonitic bivalve shells: Seasonal variations and mode of incorporation in *Arctica islandica*. *Chem. Geol.* **254**, 113–119 (2008).
249. Schöne, B. R., Zhang, Z., Raderma, P., Thébault, J., Jacob, D. R., Nunn, E. y Maures, A. Sr/Ca and Mg/Ca ratios of ontogenetically old, long-lived bivalve shells (*Arctica islandica*) and their function as paleotemperature proxies. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **302**, 52–64 (2011).
250. Toland, H., Perkins, B., Pearce, N., Keenan, F. y Leng, M. J. A study of sclerochronology by laser ablation ICP-MS. *J. Anal. At. Spectrom.* **15**, 1143–1148 (2000).
251. Lazareth, C. E., Le Cornec, F., Candaudap, F. y Freydier, R. Trace element heterogeneity along isochronous growth layers in bivalve shell: Consequences for environmental reconstruction. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **373**, 39–49 (2013).
252. Fortes, F. J., Vadillo, J. M., Stoll, H., Jiménez-Sánchez, M., Morena, A. y Laserna, J. J. Spatial distribution of paleoclimatic proxies in stalagmite slabs using laser-induced breakdown spectroscopy. *J. Anal. At. Spectrom.* **27**, 868–873 (2012).
253. Freitas, P., Clarke, L. J., Kennedy, H., Richardson, C. y Abrantes, F. Mg/Ca, Sr/Ca, and stable-isotope ($\delta^{18}\text{O}$ and $\delta^{13}\text{C}$) ratio profiles from the fan mussel *Pinna nobilis*: Seasonal records and temperature relationships. *Geochem. Geophys. Geosystems* **6**, number 4, (2005).
254. Freitas, P. S., Clarke, L. J., Kennedy, H., Richardson, C. A. y Abrantes, F. Environmental and biological controls on elemental (Mg/Ca, Sr/Ca and Mn/Ca)



- ratios in shells of the king scallop *Pecten maximus*. *Geochim. Cosmochim. Acta* **70**, 5119–5133 (2006).
255. Freitas, P. S., Clarke, L. J., Kennedy, H. A. y Richardson, C. A. Inter- and intra-specimen variability masks reliable temperature control on shell Mg/Ca ratios in laboratory- and field-cultured *Mytilus edulis* and *Pecten maximus* (bivalvia). *Biogeosciences* **5**, 1245–1258 (2008).
256. Freitas, P. S., Clarke, L. J., Kennedy, H. y Richardson, C. A. Ion microprobe assessment of the heterogeneity of Mg/Ca, Sr/Ca and Mn/Ca ratios in *Pecten maximus* and *Mytilus edulis* (bivalvia) shell calcite precipitated at constant temperature. *Biogeosciences* **6**, 1209–1227 (2009).
257. Foster, L. C., Finch, A. A., Allison, N., Andersson, C. y Clarke, L. J. Mg in aragonitic bivalve shells: Seasonal variations and mode of incorporation in *Arctica islandica*. *Chem. Geol.* **254**, 113–119 (2008).
258. Gaillard, J. M. *Aspects qualitatifs et quantitatifs de la croissance de la coquille de quelques espèces de Mollusques Prosobranches en fonction de la latitude et des conditions écologiques*. (Éditions du Muséum (impr. Lahure), 1965).
259. Regis, M. B. Étude comparée de la croissance des Monodontes (Gastéropodes Prosobranches) en Manche et le long des côtes Atlantiques et Méditerranéennes françaises. *J. Molluscan Stud.* **3**, 259–267 (1972).
260. Lorrain, A., Gillikin, D. P., Paulet, Y., Chauvard, L., Le Mercier, A., Navez, J. y André, L. Strong kinetic effects on Sr/Ca ratios in the calcitic bivalve *Pecten maximus*. *Geology* **33**, 965–968 (2005).
261. Putten, E. V., Dehairs, F., Keppens, E. y Baeyens, W. High resolution distribution of trace elements in the calcite shell layer of modern *mytilus edulis*: environmental and biological controls. *Geochim. Cosmochim. Acta* **64**, 997–1011 (2000).
262. Lutringer, A., Blamart, D., Frank, N. y Labeyrie, L. in *Cold-Water Corals and Ecosystems* 1081–1096. Springer, Berlin Heidelberg, Germany, 2005.
263. Smith, J. E., Schwarcz, H. P., Risk, M. J., McConnaughey, T. A. y Keller, N. Paleotemperatures from Deep-Sea Corals: Overcoming ‘Vital Effects’. *PALAIOS* **15**, 25–32 (2000).



264. Schröder-Ritzrau, A., Freiwald, A. y Mangini, A. in *Cold-Water Corals and Ecosystems* 157–172 Springer Berlin Heidelberg, Germany, 2005.
265. Böhm, F., Gussone, N., Eisenhauer, A., Dullo, S. y Paytan, A. Calcium isotope fractionation in modern scleractinian corals. *Geochim. Cosmochim. Acta* **70**, 4452–4462 (2006).
266. Fietzke, J. y Eisenhauer, A. Determination of temperature-dependent stable strontium isotope ($^{88}\text{Sr}/^{86}\text{Sr}$) fractionation via bracketing standard MC-ICP-MS. *Geochem. Geophys. Geosystems* **7**, Q08009 (2006).
267. Beck, J. W., Lawrance, E., Ito, E., Taylor, F., Recy, J., Rougerie, F., Joannot, P., y Henin, C. Sea-Surface Temperature from Coral Skeletal Strontium/Calcium Ratios. *Science* **257**, 644–647 (1992).
268. McCulloch, M. T., Gagan, M. K., Mortimer, G. E., Chivas, A. R. y Isdale, P. J. A High-Resolution Sr/Ca and $\delta^{18}\text{O}$ Coral Record from the Great Barrier Reef, Australia, and the 1982-1983 El Nino. *Geochim. Cosmochim. Acta U. S.* **58:12**, (1994).
269. Sinclair, D. J., Williams, B. y Risk, M. A biological origin for climate signals in corals—Trace element ‘vital effects’ are ubiquitous in Scleractinian coral skeletons. *Geophys. Res. Lett.* **33**, L17707 (2006).
270. Mitsuguchi, T., Matsumoto, E., Abe, O., Uchida, T. y Isdale, P. J. Mg/Ca Thermometry in Coral Skeletons. *Science* **274**, 961–963 (1996).
271. Bryan, S. P. y Marchitto, T. M. Mg/Ca–temperature proxy in benthic foraminifera: New calibrations from the Florida Straits and a hypothesis regarding Mg/Li. *Paleoceanography* **23**, PA2220 (2008).
272. Hathorne, E. C., Felis, T., Suzuki, A., Kawahata, H. y Cabioch, G. Lithium in the aragonite skeletons of massive Porites corals: A new tool to reconstruct tropical sea surface temperatures. *Paleoceanography* **28**, 143–152 (2013).
273. Yuta Kawakubo, Y. Y. Precise determination of Sr/Ca by laser ablation ICP-MS compared to ICP-AES and application to multi-century temperate corals. *Geochem. J.* **48**, 145–152 (2014).
274. Pandhija, S. y Rai, A. K. In situ multielemental monitoring in coral skeleton by CF-LIBS. *Appl. Phys. B* **94**, 545–552 (2009).



- 275. Motto-Ros, V., Sancey, L., Wang, X. C., Ma, Q. L., Lux, F., Bai, X. S., Panczer, G., Tillement, O. y Yu, J. Mapping nanoparticles injected into a biological tissue using laser-induced breakdown spectroscopy. *Spectrochim. Acta Part B At. Spectrosc.* **87**, 168–174 (2013).
- 276. Sancey, L., Motto-Ros, V., Busser, B., Kotb, S., Benoit, J. M., Piednoir, A., Lux, F., Tillement, O., Panczer, G. y Yu, J. Laser spectrometry for multi-elemental imaging of biological tissues. *Sci. Rep.* **4**, (2014).
- 277. Ismael Coronado, S. R. Carboniferous auloporids from the Iberian Peninsula: Palaeocology, diversity, and spatio-temporal distribution. *J. Iber. Geol.* **40**, 61–85 (2014).



Anexos:

1. Seguridad láser

La seguridad en el laboratorio es fundamental, es por eso adecuado introducir las principales normas y precauciones a la hora de trabajar con o cerca de la radiación láser. La última revisión de la seguridad del láser se basa en la normativa europea “New laser classification (UNE EN 60825-1/A2: 2002)”, que actualiza a la NTP-261 de 1991. Dependiendo de las características técnicas de cada láser como su sistema de bombeo, medio activo y cavidad óptica, los láser pueden emitir en un amplio rango de longitudes de onda, con diferentes potencias de salida y pudiendo ser de emisión continua o pulsada. Por tanto, debido a su heterogeneidad no es posible agrupar todos los láseres dentro de un mismo grupo de seguridad y es necesario realizar una clasificación para adecuar las normas de seguridad al tipo de láser. El uso o aplicación de la radiación láser condiciona el tiempo de exposición, siendo este un factor clave para determinar el riesgo. Como consecuencia de nuevos sistemas láser la antigua clasificación en clases 1, 2, 3B y 4 ha quedado obsoleta. La modificación solo afecta a los láseres de antigua tipo 3B, que ahora se dividen en tres nuevas clases con potencias intermedias como las 1M, 2M y 3R.

Los láseres de clase 1 son seguros en todas las condiciones de utilización razonablemente previsibles, incluyendo el uso de instrumentos ópticos en visión directa. No causan daño durante una exposición directa al haz láser por un periodo de 8 horas, independiente de la longitud de onda.



irreversible en ojos y piel. La radiación láser puede afectar al organismo a través de tres procesos principalmente: térmicos, acústicos y fotoquímicos.

Los procesos térmicos se producen como consecuencia a un incremento de la temperatura en una zona debido a la absorción de radiación. El daño producido depende de factores como la duración de la exposición, la longitud de onda incidente, la energía del haz y la superficie y tipo de tejido irradiado. Los efectos acústicos son en general menos importantes y se producen como resultado de recibir la onda de choque generada por el pulso láser y ser propagada a través del tejido produciendo un daño en el mismo.

Los efectos fotoquímicos se producen cuando los fotones interactúan con las células del tejido. Esta interacción puede producir un cambio en la química de las células afectadas, siendo la longitud de onda incidente el parámetro fundamental para evaluar el daño producido.

En función de la longitud de onda y la zona afectada (ojos o piel) se puede simplificar los efectos en la siguiente Tabla 8.1.

Tabla 8.1: Efecto de la radiación láser sobre la salud humana

Región espectral de emisión	Ojo	Piel
UVC: 200-280nm	Fotoqueratitis	Quemadura similar a la del sol, cáncer de piel, envejecimiento acelerado de la piel
UVB: 280-315nm	Fotoqueratitis	Incremento de la pigmentación de la piel
UVA: 325-400nm	Catarata fotoquímica	Pigmentación y quemadura cutánea
Visible: 400-780nm	Lesión retinal de origen térmico y fotoquímico	Reacciones de fotosensibilización, quemadura cutánea
IR A: 780-1400nm	Catarata y quemadura de retina	Quemadura cutánea
IR B: 1.4-3.0μm	Catarata, quemadura de córnea, inflamación del humor acuoso	Quemadura cutánea
IR C: 3.0-1000μm	Quemadura de Cornea	Quemadura cutánea

Además del daño producido por la propia radiación láser, es importante considerar los efectos sobre la salud que puede provocar en alguna aplicación como por ejemplo en la ablación láser. Durante la ablación laser la pluma generada provoca una pequeña “nube” de contaminantes debido a la interacción del láser con el material de análisis. Esta pluma puede contener humos metálicos, partículas en suspensión y aerosoles con material biológico. La pluma generada es de pequeño tamaño y la cantidad de contaminantes a los que se expone el personal es bajo, sin embargo una exposición prolongada puede producir efectos nocivos sobre la salud, por tanto para evitar o disminuir el riesgo es aconsejable trabajar con una ventilación local en la zona de ablación o usar un sistema de extracción propio del equipo. Por último en casos donde el material a analizar sea susceptible de producir un riesgo para la salud será obligatorio

el uso de mascarar de protección además de las medidas de seguridad expuestas con anterioridad.

2. Software e interfaces de usuario

Durante el desarrollo de este tesis se han programado en Matlab varias rutinas y software útiles. En primer lugar se realizó una interfaz de usuario para la adquisición y pre procesamiento espectral. La Fig. 8.1 muestra la interfaz de usuario generada, en la parte izquierda se encuentra el panel de carga de datos, los paneles de texto de arriba abajo muestras: Nombre del archivo, nombre de la librería a generar, número del primer espectro y por último el número del último espectro. Por otro lado, a la derecha aparece un gráfico donde se muestran los espectros procesados con el objetivo de detectar posibles fallos en el programa de forma rápida y visual.

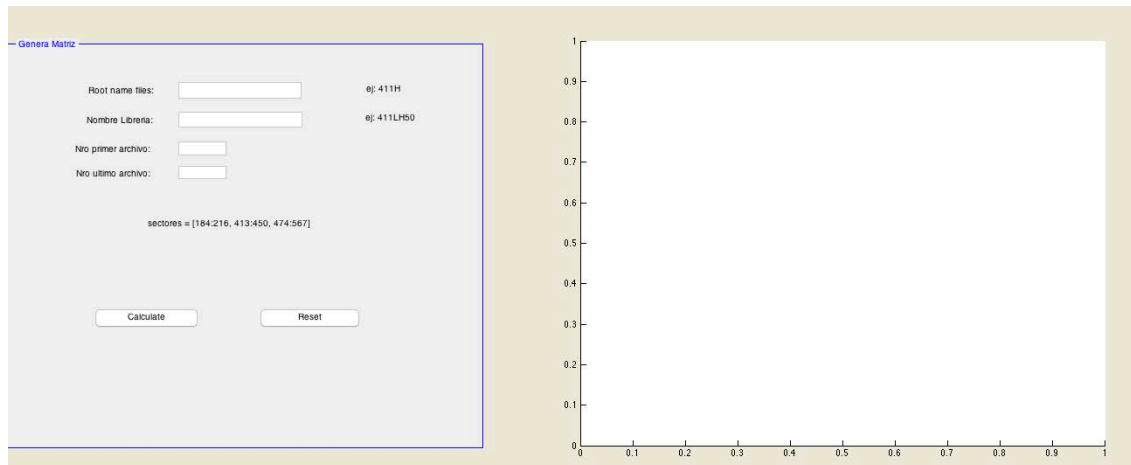


Figura 8.1: Interfaz, gráfica procesamiento de espectros

Se programó y realizó una interfaz para la realización del análisis de componentes principales en Matlab. Previamente se realizaba en un software comercial, The Unscramble (CAMO), sin embargo su funcionamiento es lento y la importación de espectros tediosa. Se desarrolló por tanto un programa dentro de Matlab que permita la

realización de un análisis de componentes principales como primer paso en cualquier tratamiento de datos, con el objetivo de observar posibles tendencias y eliminar datos anómalos. La Fig. 8.2 muestra la interfaz generada. Como se puede observar se intentó simplificar al máximo para que cualquier usuario pueda llevar a cabo un análisis PCA aunque no se tengan conocimiento del mismo. Al igual que la interfaz anterior, esta también cuenta con dos partes, a la izquierda se cargan las matrices (objetos por longitudes de onda), se selecciona el número de espectros de cada y se ejecuta el programa haciendo *click* en “Empezar PCA”. En el gráfico de la derecha aparece el análisis de PCA obtenido marcando con diferentes colores cada una de las clases. Se intentó simplificar al máxima y por eso, para simplificar, solo permite introducir el mismo número de objetos para todas las clases. Haciendo una pequeña modificación es posible realizar el análisis PCA con diferente número de objetos por muestra.

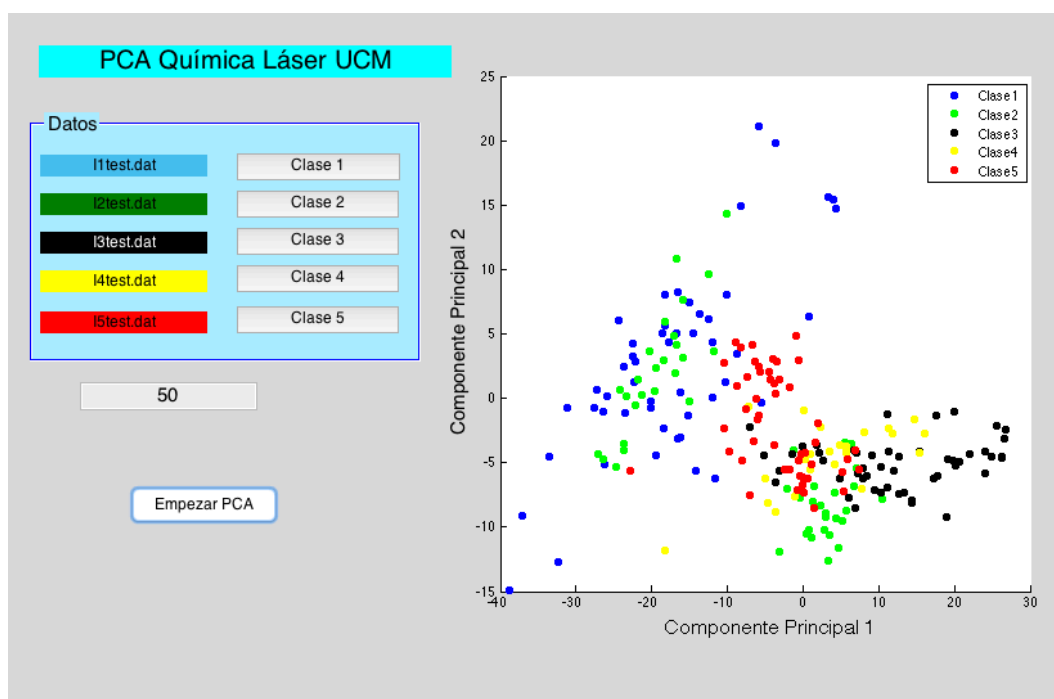


Figura 8.2: Interfaz gráfica análisis de componentes principales (PCA)

En cuanto a las redes neuronales se desarrollaron dos interfaces, una para cada etapa del proceso. Dichas interfaces se realizaron con el mismo objetivo que en el caso del PCA, ser lo más sencillas posibles y que cualquier usuario sin formación específica sea capaz de utilizarlas. Primero para el entrenamiento de la red neuronal se diseñó la interfaz mostrada en la Fig. 8.3, en este caso se trata de una red binaria, donde solo se puede entrenar con dos muestras diferentes. Estas han de ser cargadas haciendo *click* en “*Select Reference 1*” y “*Select Reference 2*”, respectivamente. Posteriormente hemos de seleccionar el número de espectros de cada librería de referencia y por último el número de neuronas de la capa oculta que deseamos usar para el entrenamiento. Haciendo *click* en “*Training*” comienza el entrenamiento y se abre una nueva pantalla que permite observar el proceso de entrenamiento. En nuestra interfaz se aparece un resumen del proceso que permite al usuario observar el comportamiento de la red por si es necesario volver a entrenar el modelo y modificar algún parámetro de la misma.

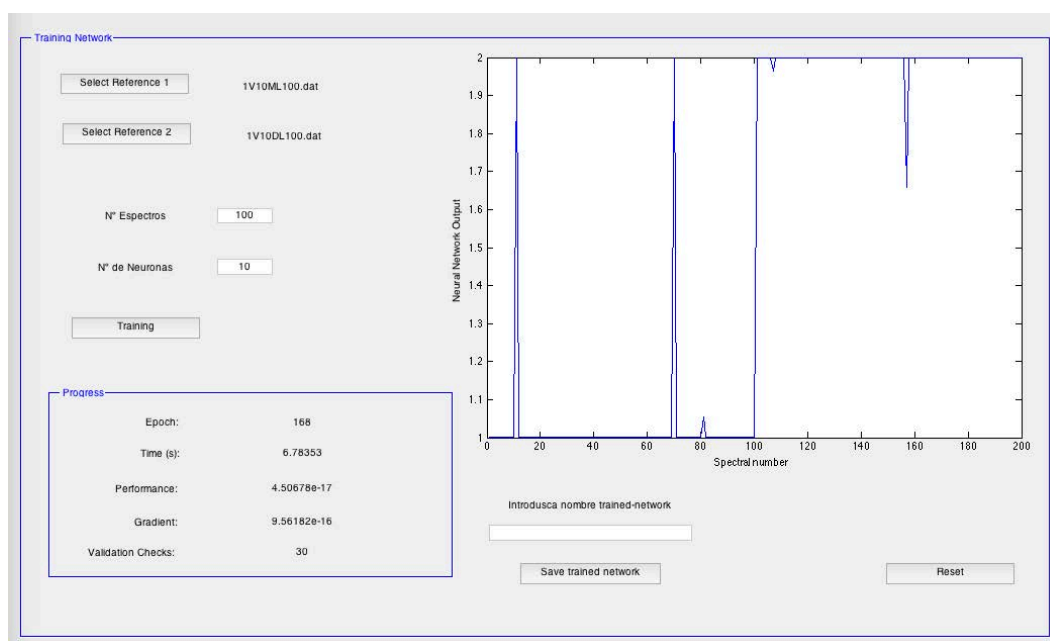


Figura: 8.3 Interfaz gráfica entrenamiento de la Red Neuronal



Como segunda parte del proceso de análisis por redes neuronales se realizó una interfaz para el test. La figura [Fig. 8.4](#) muestra la interfaz desarrollada. Para su realización basta con cargar la librería que queremos testar, seleccionar el modelo de Red Neuronal, introducir el número de espectros que tiene nuestra librería y hacer *click* en “*Test File*”. Al hacerlo aparecen dos ventanas nuevas, en primer lugar, debajo de la carga de datos aparece un resumen de los resultados de clasificación, nos indica el índice de coincidencia (*Match Index*) a la Clase 1, Clase 2 y espectros clasificados como desconocidos. A la derecha se puede observar el resultado gráficamente. En el ejemplo se introdujo una muestra de vino de Ribera de Duero a un entrenamiento generado con Ribera de Duero y Vino de la Mancha, siendo la clase 1 y 2, respectivamente. Como se puede observar en la figura todos los espectros son clasificados como pertenecientes a la clase 1. Este tipo de gráficos es de gran utilidad ya que permite observar cual o cuales espectros están incorrectamente clasificados y así detectar posibles problemas de medida.

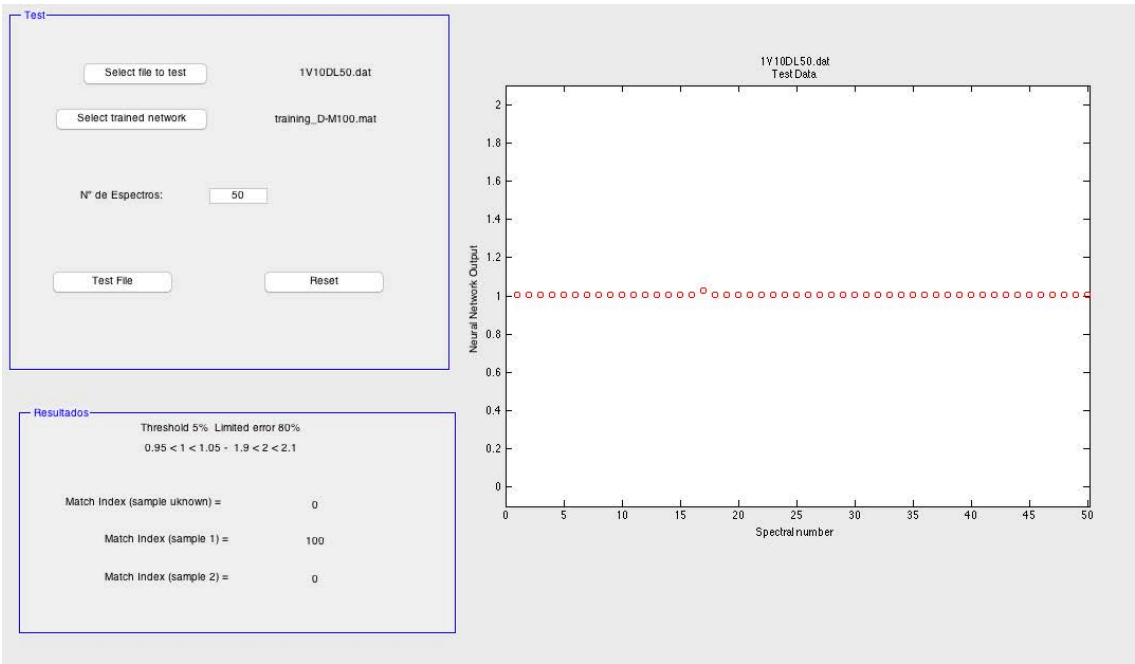


Figura 8.4: Interfaz gráfica del test de la Red Neuronal



3. Resultados de la estancia en la Universidad Comenius

Discrimination of Copper Alloys with Archaeological Interest Using LIBS and Chemometric Methods

S. Moncayo,^{1,2} M. Kociánová,¹ J. Hulík,¹ J. Plavčan,¹ M. Hornáčková,¹ M. Suchonňová,¹ P. Veis,¹ J. O. Cáceres²

¹ Department of Experimental Physics, Faculty of Mathematics, Physics and Informatics, Comenius University, Mlynská dolina F2, 842 48 Bratislava, Slovakia.

² Department of Analytical Chemistry, Faculty of Chemistry, Complutense University, 28040 Madrid, Spain.

Abstract:

In this work, the discrimination between six copper alloys with archaeological interest has been carried out applying Laser Induced Breakdown Spectroscopy (LIBS) and two different chemometric methods; Principal Component Analysis (PCA) and Soft Independent Modelling of Class Analogy (SIMCA). Plasma optical emission was analyzed in the range of 200–1000 nm with an Echelle spectrometer and an ICCD camera. These copper alloys contained different amounts of Sn and other minority elements such as Pb, Al, Zn, Ni, Fe and Mn. The results show the potential applicability of the methodology in the classification of bronze alloys with a high accuracy and robustness.

Introduction

Metallic artifacts are one of the most rewarding objects that can be found in



archaeological exploration. The composition of the artifacts is related to its function and the necessity to assess their chemical composition is always a priority. The characterization of metal artifacts is always the first step in proposing a conservation treatment or preventive care measures. There are different techniques available to obtain the elemental composition of metal artifacts such as X-ray fluorescence and diffraction spectroscopy, Raman spectroscopy, Fourier transform infrared spectroscopy, scanning electron microscopy and inductively-coupled-plasma atomic-emission spectrometry among others.¹⁻³ All these techniques are able to achieve high accuracy and precision. However, in some cases, the sampling step involves the need to take a portion of the sample and transfer it to the laboratory for the analysis and also involves the dissolution of the sample which destroys the sample. Moreover, in several sort of samples, due to the size and the condition of samples this approach is not always possible. The advantages of using LIBS in archaeology have been shown in many investigations.^{4,5} The possibility to carry out analysis in field, its quasi non-destructive performance and their little or no sample preparation are the most important advantages in archaeological applications, reducing the analysis time and avoiding possible stages where the samples could be damaged. LIBS has been used in cleaning the surface of samples of cultural heritage, as well as in identification and restoration of artifacts of cultural interest. Several studies have been published recently using LIBS for the characterization of archaeological bronze.^{6,7}

To achieve the sample identification, LIBS produces a “spectral fingerprint” of the sample under analysis and the posterior analysis of the spectral data by chemometric methods allow to obtain relevant information of samples, enabling their classification.

In this study LIBS has been used in combination with Principal Component Analysis (PCA) and Soft Independent Modelling of Class Analogies (SIMCA) to develop a quick and cost-effective methodology for the classification of archaeological bronze samples.

Experimental LIBS Set up

A simplified schematic diagram of the LIBS set-up has been shown in Figure 1. The experiment was conducted using a Q-switched Nd:YAG laser operating at 532 nm producing pulses of 4 ns, maximum repetition rate of 10 Hz and maximum energy of 165 mJ per pulse. A scheme with the set-up is shown in the Fig.1. Laser beam was focused on the sample using a 40 mm focal length quartz lens generating a spot of 0.1 mm of diameter. Plasma emission was collected using a mirror system and focused into the optical fiber directly coupled to the entrance slit of an echelle spectrometer (Mechelle ME 5000, Andor Technology) providing a recording of a wide wavelength range (200–950 nm) in one acquisition. An intensified charge coupled device (ICCD) camera (iStar, Andor Technology) was used as detector. The gate width was set to 500 ns and the delay time after the laser shot was set to 1 μ s. For wavelength calibration, mercury-argon calibration lamp (HG-1, Ocean Optics) was used.

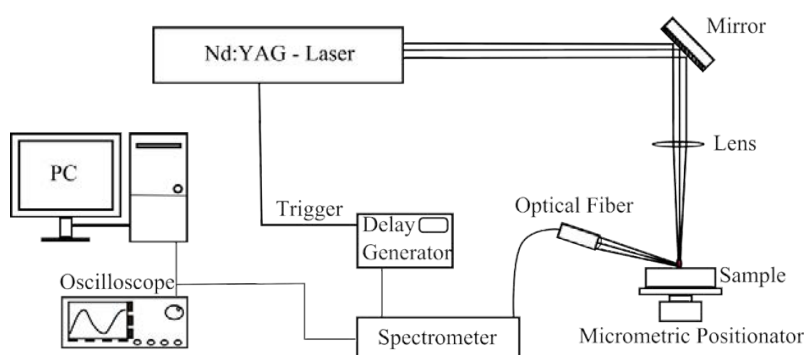


Figure 1: Scheme of the LIBS experimental set-up.

Description of chemometric methods used

The Principal Component Analysis (PCA) is a powerful statistical technique for finding patterns in data with high dimensionality, such as LIBS spectra. It is a multivariate lineal non-supervised classification method that is able to make lineal combinations of the original variables reducing the number of dimensions and making possible the visualization of the data in graphical representations helping to understand the raw data.⁸

Otherwise SIMCA is the most used supervised chemometric method for classification purpose. It is able to classify unknown objects to their probably class membership. In SIMCA analysis, a PCA model is performed for each class. Then a cross-validation is performed until the model can determine each data in their class membership. To classify a new object, an unknown sample, the residual variance is compared to the average of the modelled variance for each class. This residual parameter is called Mahalanobis distance. This distance determines to which model the new object belongs to and makes the classification accordingly. If all the distances calculated are outside to a fixed value, the new sample is unclassified (unknown).

Alloy Samples

A total of 6 different alloys were used in the study. The samples covered a range of binary (Sn- Cu), ternary (Sn-Cu-Pb) and quaternary (Sn-Cu-Zn-Pb, Cu-Al-Ni-Fe or Cu-Al-Fe-Mn) reference alloys, characterized by a chemical composition and metallurgical features similar to found in archeological artifacts in Slovakia. A complete chemical composition in % w/w of the samples is shown in Table 1. For each alloy a total of 10 spectra were collected where each spectrum was an accumulation of 10 laser shots.

Table 1: Chemical composition in % w/w of the samples

References	Alloy 1	Alloy 2	Alloy 3	Alloy 4	Alloy 5	Alloy 6
Chemical Composition (%)						
Cu	81.87	87.51	79.204	91.93	86.076	78.8
Sn	5.6	11.17	0.002	7.9	0.009	9.57
Pb	7.4	0.6	0.003	0.02	0.001	9.05
Zn	4.56	0.42	0.1	0.01	0.07	1
Fe	0.1	0.04	4.7	0.01	2.59	0.04
Ni	0.47	0.26	4.98	0.01	0.111	1.28
P	*	*	*	0.12	*	0.07
Al	*	*	10.7	*	9.3	*
Mn	*	*	0.19	*	1.82	0.001
Si	*	*	0.001	*	0.023	*
Sb	*	*	*	*	*	0.14
S	*	*	*	*	*	0.0001

Results

Figure 2a and 2b show the LIBS spectra for alloy 1 and 3. The spectral lines assignation was done based on NIST atomic spectra database [9]. In alloy 1 emission lines from Pb can be seen in addition to Cu and Sn whereas no signals of Al were observed. In case of alloy 3 Al, Mn and Ni in addition to Cu lines were visible.

The first procedure performed was PCA. In this case all alloy spectra were used, however, the variables were selected in several wavelength ranges where the most intense spectral lines were present, reducing the number of variables from initially 20,000 to 900. With PCA analysis it is possible to make an initial data exploratory analysis observing an overview of the raw data in simple graphs. The obtained score plot, representing the first and second principal component arranging an 82% of the total variance is shown in Figure 3. It allows divide the data in three different groups of samples. The first group represented in the second quadrant involved alloy 2 and 4,

which were characterized for containing only Cu and Sn in their composition, whereas other elements were only in

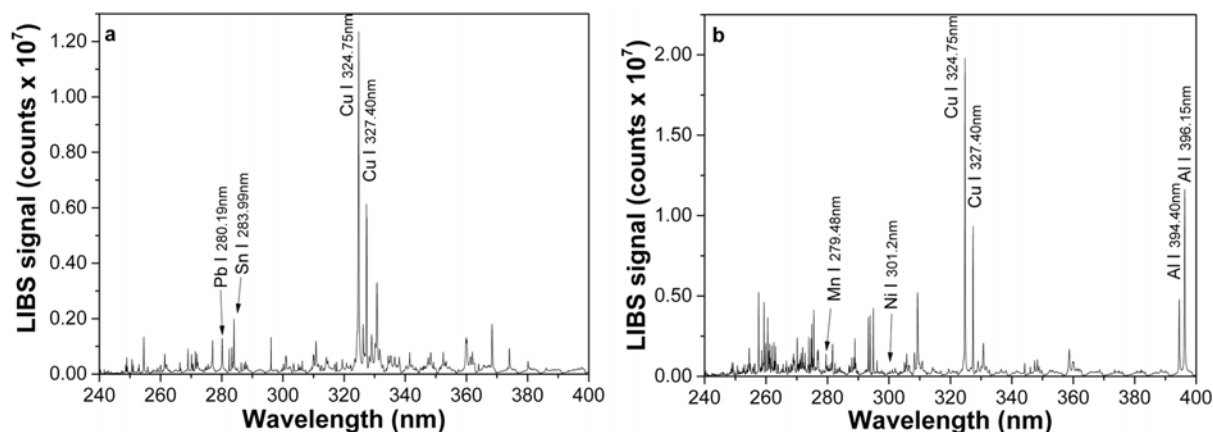


Figure 2: Accumulate spectrum of (a) alloy 1 and (b) alloy 3.

trace amounts. In the second group, alloys with Pb were separated from the remaining alloy samples. The third group was separated from other samples because of the presence of Al in the alloys. It was not possible to see other tendencies due to other elements like Mn or Ni because of their low concentration in the alloys and their weak LIBS signals. Thus, with PCA it was possible to discriminate between bronze alloys based on the presence of Al, Pb or only Sn-Cu. However, it was not possible to completely discriminate between the six alloys and therefore a more complex chemometric method, such as SIMCA, was necessary to obtain the corresponding classification.

SIMCA was applied to the classification of the spectra for all alloys. The data set was split into a calibration set by choosing the first 7 spectra for each alloy, whereas the test set was created by the remaining 3 spectra. An individual PCA analysis was carried out for each alloy, resulting in six independent PCA models. It can be seen in the Figure 3 that the minor distances correspond to the alloys 2–4, 1–6, and 3–5 respectively. For the

prediction of “unknown” spectra, test set, each PCA model was applied with their corresponding suitable number of principal components for the analysis, chosen by cross-validation. The number of principal components in all cases was between three and five. Once SIMCA defined a sub-space for each alloy, it is possible to predict the class membership of the test set (unknown). Each spectrum is projected into the corresponding sub-space and the distances between the spectrum and all classes were assessed. As can be seen from the results in the Table 2, all spectra except the eighth spectrum for alloy 5 were correctly classified with no unclassified spectra and moreover no spectra incorrectly classified to another class. In the case of the eighth spectrum for the alloy 5, SIMCA was incapable to classify in its corresponding alloy and the result was unclassified.

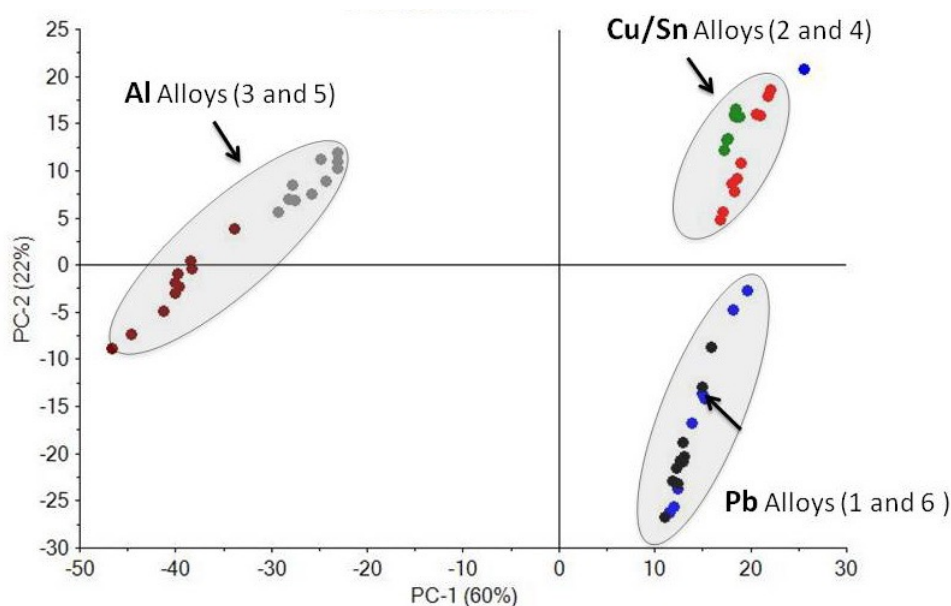


Figure 3: PCA analysis (scores PC2 vs scores PC1) for all alloy spectra used.

Table 2: Classification Results of alloys samples using SIMCA.

Validation		Predicted Group Membership					
Actual Alloy	Sample ID	Alloy 1	Alloy 2	Alloy 3	Alloy 4	Alloy 5	Alloy 6
CuSn7Zn4Pb7	Alloy 1-S8	1	0	0	0	0	0
	Alloy 1-S9	1	0	0	0	0	0
	Alloy 1-S10	1	0	0	0	0	0
CuSn12	Alloy 3-S8	0	1	0	0	0	0
	Alloy 3-S9	0	1	0	0	0	0
	Alloy 3-S10	0	1	0	0	0	0
CuAl10Ni5Fe7	Alloy 4-S8	0	0	1	0	0	0
	Alloy 4-S9	0	0	1	0	0	0
	Alloy 4-S10	0	0	1	0	0	0
CuSn8	Alloy 5-S8	0	0	0	0	0	0
	Alloy 5-S9	0	0	0	1	0	0
	Alloy 5-S10	0	0	0	1	0	0
CuAl10Fe3Mn2	Alloy 6-S8	0	0	0	0	1	0
	Alloy 6-S9	0	0	0	0	1	0
	Alloy 6-S10	0	0	0	0	1	0
CuSn10Pb10	Alloy 7-S8	0	0	0	0	0	1
	Alloy 7-S9	0	0	0	0	0	1
	Alloy 7-S10	0	0	0	0	0	1

Conclusions

A set of 60 spectra for 6 different alloys of archaeological references was classified using chemometric methods. PCA allowed to arrange the alloys in three groups according to the main elements doped with Cu such as Al and Pb. However, PCA did not achieve a complete separation into six groups. Using SIMCA as prediction method, all LIBS spectra were correctly classified to their corresponding class membership with the exception of one spectrum for the alloy 5 that was unclassified. Therefore, the combination of LIBS and Chemometric methods has shown its potential applicability as a methodology to carry out a fast and cost-effective classification of bronze alloys that would help in setting the corresponding conservation procedures.

In order to get a complete classification of the spectra, a complementary analysis with



more complex chemometric method such as, Partial Least Square Discriminant Analysis (PLS-DA) or the use of non-linear methods as Neural Network should be carried out and assessed.

References of this work

- [1] Enrico Ciliberto G S. Modern Analytical Methods in Art and Archeology. First Edition. New York: Wiley, 2000.
- [2] Beck L, Bosonnet S, Réveillon S, Eliot D, Pilon F. Silver surface enrichment of silver-copper alloys: a limitation for the analysis of ancient silver coins by surface techniques. *Nucl Instrum Methods Phys Res Sect B Beam Interact Mater At*, 226 (1–2), 153–62, 2004.
- [3] Mohamed W, Darweesh S. Ancient Egyptian Black-Patinated Copper Alloys. *Archaeometry*, 54 (1), 175– 192, 2012.
- [4] Fortes F J, Cortés M, Simón M D, Cabalín L M, Laserna J J. Chronocultural sorting of archaeological bronze objects using laser-induced breakdown spectrometry. *Anal Chim Acta*, 554 (1–2), 136–43, 2005.
- [5] Elhassan A, Giakoumaki A, Anglos D, Ingo GM, Robbiola L, Harith MA. Nanosecond and femtosecond Laser Induced Breakdown Spectroscopic analysis of bronze alloys. *Spectrochim Acta Part B At Spectrosc*, 63 (4), 504–511, 2008.
- [6] Colao F, Fantoni R, Lazic V, Spizzichino V. Laser induced breakdown spectroscopy for semi quantitative and quantitative analyses of artworks: application on multi layered ceramics and copper based alloys. *Spectrochim Acta Part B At Spectrosc*, 57 (7), 1219–



1234, 2002.

[7] Ferretti M, Cristoforetti G, Legnaioli S, Palleschi V, Salvetti A, Tognoni E, et al. In situ study of the Porticello Bronzes by portable X-ray fluorescence and laser-induced breakdown spectroscopy. *Spectrochim Acta Part B At Spectrosc*, 62 (12), 1512–1518, 2007.

[8] Bro R, Smilde AK. Principal component analysis. *Anal Methods*, 6 (9), 2812-31, 2010.

[9] NIST. Atomic Spectroscopy Information. Cited 2014 May 26. <http://www.nist.gov/atomicspectroscopy>.

Application of Laser-Induced Breakdown Spectroscopy (LIBS) and Neural Networks to Olive Oils Analysis

Jorge O. Caceres,^{a,*} Samuel Moncayo,^a Juan D. Rosales,^a Francisco Javier Manuel de Villena,^a Fernando C. Alvira,^b Gabriel M. Bilmes^{b,c}

^a Departamento de Química Analítica, Facultad de Ciencias Químicas, Universidad Complutense, 28040 Madrid, Spain

^b Centro de Investigaciones Ópticas (CONICET La Plata-CIC) CC 3, 1897, Gonnet, La Plata, Buenos Aires, Argentina

^c Facultad de Ingeniería Universidad Nacional de La Plata, La Plata, Buenos Aires, Argentina

The adulteration and traceability of olive oils are serious problems in the olive oil industry. In this work, a method based on laser-induced breakdown spectroscopy (LIBS) and neural networks (NNs) has been developed and applied to the identification, quality control, traceability, and adulteration detection of extra virgin olive oils. Instant identification of the samples is achieved using a spectral library, which was obtained by analysis of representative samples using a single laser pulse and treatment by NNs. The samples used in this study belong to four countries. The study also included different regions of each country. The results obtained allow the identification of the oils tested with a certainty of more than 95%. Single-shot measurements were enough for clear identification of the samples. The method can be developed for automatic real-time, fast, reliable, and robust measurements, and the system can be packed into portable form for non-specialist users.

Index Headings: Laser-induced breakdown spectroscopy; LIBS; Neural networks; Olive oils; Edible oils.

INTRODUCTION

Control of adulteration, together with authentication and traceability, is an important problem in the quality control of olive oils. In particular, adulteration with lower-price oils is one of the most difficult situations to detect and is a serious problem in the olive oil industry, because of the similar composition of the adulterant oils used.

The high price in the market and the known health benefits attributed to olive oil^{1,2} make this product a target for different types of fraud, such as adulteration by blending with other oils of different sources. This is especially true for extra virgin olive oil, which fetches much higher prices in international markets than any other vegetable oil.

Besides the intentional addition of oils from different sources to extra virgin olive oil during the production process, there is also a possibility of unwanted contamination. Different analytical tests are described in the literature to detect small additions of low-quality olive oil,^{3,4} hazelnut oils,⁵ and seed oils⁶ in the extra virgin product.

The use of techniques such as gas chromatography,⁷ reverse-phase high-performance liquid chromatography,^{8,9} gas chromatography coupled to mass spectrometry,^{10,11} and nuclear magnetic resonance (NMR)^{12–14} provides authentication and detection of adulterants in quantities less than 1% of foreign oils in extra virgin olive oil. Although these analytical techniques are widely used in authentication, quality control, and adulteration detection of edible oils, they have the

drawback of being expensive or involve a time-consuming derivatization step.

Vibrational spectroscopy such as Fourier transform infrared spectroscopy,^{15,16} mid-infrared spectroscopy,¹⁷ and Raman techniques^{18,19} has been used to detect adulterant in olive oils for many years.^{20,21} The use of chemometric methods for data analysis has greatly improved the results obtained using these techniques.^{17,18,22} In combination with principal component analysis (PCA), partial least squares, mode recognition, or neural network (NN) techniques, Raman spectroscopy has been successfully used to identify and quantify adulteration of olive oils with corn oils, soybean oils, olive pomace oils, hazelnut oils, etc.^{22–24} Although many of the spectral features are similar, there are nevertheless sufficiently large differences in the Raman spectra for unambiguous identification.²⁵ However, Raman spectroscopy is not routinely used in many situations because of high background scattering and time-consuming sample alignment procedures.^{26,27}

Laser induced breakdown spectroscopy (LIBS) has been a subject of research for the past few decades because of its unique features and the wide variety of applications in various fields.^{28,29} In recent years, LIBS has become a powerful analytical tool because of its ability to carry out a rapid qualitative and quantitative analysis of different samples.^{30,31} LIBS analyzes a sample by direct measurement of the atomic emission of the elements from a laser-induced plasma generated by the ablation of the sample, providing an immediate spectral fingerprint that is representative of its elemental composition.³² Moreover, the technique requires little or no sample preparation. Although there is a loss of molecular information in plasma, this technique has provided excellent results for the identification of many polymer organic compounds,³³ explosives,³⁴ and bacteria.³² The intensities of C, H, N, and O lines and the C/H, C/O, and C/N intensity ratios have been used for the classification of organic compounds.³⁴

The aim of this paper is to use a simple and direct method to identify and discriminate edible oils based on LIBS and NNs. The method developed relies on the instantaneous identification of the oil sample using a unique feature of LIBS, which is its ability to generate a spectral “fingerprint” of the sample because of the nature of the emission spectra, representative of the main elements that constituted the sample. Thus, LIBS provides a unique spectrum, corresponding only to the sample under analysis. Using a correlation procedure, the developed LIBS-NN system, can be trained to recognize spectra from different samples, evaluating the similarity of unknown spectra against a spectral library of classified samples.

Several chemometric methods used to process LIBS data have been evaluated by different research groups, such as PCA,

Received 15 November 2012; accepted 29 April 2013.

* Author to whom correspondence should be sent. E-mail: jcaceres@ucm.es.
DOI: 10.1366/12-06916



Technical Note

Discrimination of human bodies from bones and teeth remains by Laser Induced Breakdown Spectroscopy and Neural Networks



S. Moncayo, S. Manzoor, T. Ugidos, F. Navarro-Villoslada, J.O. Cáceres*

Departamento de Química Analítica, Facultad de Ciencias Químicas, Universidad Complutense, 28040 Madrid, Spain

ARTICLE INFO

Article history:

Received 7 February 2014

Accepted 15 July 2014

Available online 23 July 2014

Keywords:

Laser Induced Breakdown Spectroscopy

Neural Networks

Human remains

Bone

Teeth

ABSTRACT

A fast and minimally destructive method based on Laser Induced Breakdown Spectroscopy (LIBS) and Neural Networks (NN) has been developed and applied to the classification and discrimination of human bones and teeth fragments. The methodology can be useful in Disaster Victim Identification (DVI) tasks. The elemental compositions of bone and teeth samples provided enough information to achieve a correct discrimination and reassembling of different human remains. Individuals were classified with spectral correlation higher than 95%, regardless of the type of bone or tooth sample analyzed. No false positive or false negative was observed, demonstrating the high robustness and accuracy of the proposed methodology.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

The proper identification of human bodies is important from the legal and administrative point of view. However, this identification issue is sometimes extremely difficult, particularly in mass disaster situations or genocide crimes [1,2]. Adequate human body identification management goes through strict methodology and protocols. This requires adopting and applying standard operating protocols such as the Interpol Disaster Victim Identification (DVI) Standing Committee guidelines [3]. When the number of remains is high, such as in natural disasters, accidents and mass graves, finding the identity of an individual is required. The necessity to achieve an accurate identification is not only important from the humanitarian approach but is also important in lawsuits. The identification process involves a multidisciplinary group of experts and techniques, i.e. fingerprint analysis, forensic pathology, forensic odontology and DNA analysis. When the bodies are non-decomposed the methods used in DVI include physiognomic data analysis such as scars or marks, personal effects, matching of fingerprints and dentition pattern (provided premortem records are available) [2,4,5]. In many situations, these methods cannot be used due to either extensive putrefaction and destruction of the remains or unavailability of appropriate medical or dental records. Common methods for human identification are not sufficient in approximately

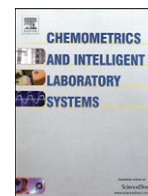
42% of cases, because of decomposition of the body and unavailability of pre-mortem data and therefore DNA identification was requested [6]. However, the unavailability of DNA from relatives, the complications in evidence collection and the degradation and contamination of DNA extracted from bone and teeth samples can hinder the identification process [7]. Moreover, in case of huge number of remains, even identification by DNA analysis becomes difficult, requiring more time and economical resources.

Therefore, a new methodology that provides simple, direct and cost-effective analysis is needed in forensic science. In recent years, Laser Induced Breakdown Spectroscopy (LIBS) has become a powerful analytical tool because of its ability to carry out a rapid qualitative and quantitative analysis of different samples, able to provide real time spectral fingerprint of the elemental composition of the sample [8–10]. The possibility to combine LIBS with chemometric methods allows extracting significant information to classify different samples, such as providing an acceptable classification of commercial pharmaceutical tablets using PCA and SIMCA [11].

Human bones and teeth are the typical remains that can be found in forensic scenes due to their resistance to degradation and, hence, are useful as evidences in anthropology, archeology and forensic science [12,13]. These remains are useful to classify and discriminate individuals in crime scenes, accidents or mass burial sites. Although differences in the elemental composition of these samples between individuals are negligible, some trace elements can be related to diet and environment [7], providing variations in the LIBS spectra. These differences allow the achievement of the discrimination when a chemometric analysis is done.

* Corresponding author.

E-mail address: jcaceres@ucm.es (J.O. Cáceres).



Evaluation of supervised chemometric methods for sample classification by Laser Induced Breakdown Spectroscopy

S. Moncayo, S. Manzoor, F. Navarro-Villoslada, J.O. Cáceres *

Department of Analytical Chemistry, Faculty of Chemistry, Complutense University, 28040 Madrid, Spain

ARTICLE INFO

Article history:

Received 9 February 2015

Received in revised form 14 May 2015

Accepted 10 June 2015

Available online 17 June 2015

Keywords:

Laser Induced Breakdown Spectroscopy

Chemometrics

Supervised methods

Bone remains

ABSTRACT

In this work seven supervised chemometric methods have been evaluated in a real world application for the classification of human bone remains with similar elemental composition based on Laser Induced Breakdown Spectroscopy (LIBS) measurements. Bone samples belonging to five individuals were obtained from a local cemetery, exposed to uncontrolled conditions. LIBS data were processed with different linear and non-linear supervised chemometric approaches. The performance of each chemometric model was assessed by three validation procedures taking into account their sensitivity (internal validation), generalization ability and robustness (independent external validation). The accuracy of each method increased in the following order: 42% for Linear Discriminant Analysis (LDA), 48% for Classification and Regression Tree (CART), 56% for Support Vector Machines (SVM), 58% for Soft Independent Modeling of Class Analogy (SIMCA), 58% for Partial least Squares–Discriminant Analysis (PLS-DA), 66% for Binary Logistic Regression (BLR) and 100% for Artificial Neural Networks (NN). The results showed that NN outperforms in terms of sensitivity, generalization ability and robustness; whereas SIMCA, PLS-DA, LDA, CART, Logistic Regression and SVM did not show significant accuracy to discriminate the bone samples with a high degree of similarity.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Chemometrics has been established as a mature methodology for converting huge amounts of measurements into significant and useful information in a multitude of disciplines and applications, for instance, signal processing, experimental design, optimization, data mining, multivariate calibration and classification [1–4]. In order to obtain maximum information from the raw data, a variety of statistical and computational algorithms have been applied. The selection of a chemometric method depends on the analytical application and the required information. Regarding classification methods, two possible procedures can be applied. Firstly, unsupervised methods, where no prior classification model is assumed over data collected in a matrix (objects). Secondly, supervised methods, which are defined by two data sets, objects (input) and classes (target). These methods attempt to discover a relationship between objects and classes, referred to as a model, which represents a set of features that define the classification process. The membership of new objects (unknowns for the models) is predicted on the basis of their similarity to a certain class in the model.

The mathematical algorithms used by these methods are based on two different approaches, linear and non-linear. Linear chemometric

methods use linear combination of variables to estimate the classification model; whereas non-linear methods apply more complex algorithms increasing the sophistication and the classification capacity of the method.

In Laser Induced Breakdown Spectroscopy (LIBS) the interaction of a high energy focused laser pulse with the sample surface is capable of producing the rupture of chemical bonds and the excitation of atoms and ions, in a process called ablation [5–7]. The optical emission of the plasma contains information about the sample and its spectroscopic study provides useful data for qualitative and quantitative applications [8,9]. LIBS technique has a broad scope due to the possibility of analyzing solids, liquids or gases, being able to acquire quickly a large amount of data. LIBS provides a spectral fingerprint of the elemental composition of the sample, which allows to obtain classification results in real time [8,10].

A wide range of chemometric approaches are available to carry out classification and discrimination processes [11–13]. Nevertheless, the selection of the appropriate method in each case is not straightforward and the raw data must be carefully studied in order to achieve a satisfactory classification result.

An appropriate chemometric method must be able to perform a successful classification when all samples included in the training are tested and correctly classified (high sensitivity), as well as unknown samples of the same class not included in the training step are also correctly classified (high generalization ability). Furthermore, the model must be able to classify a sample that does not belong to any

* Corresponding author. Tel.: +34 91 3944322; fax: +34 91 3944329.
E-mail address: jcaceres@ucm.es (J.O. Cáceres).

PAPER

CrossMark
click for updatesCite this: *J. Anal. At. Spectrom.*, 2015, 30, 1913

Mg/Ca ratios measured by laser induced breakdown spectroscopy (LIBS): a new approach to decipher environmental conditions

A. García-Escárcaga,^{†a} S. Moncayo,^{†b} I. Gutiérrez-Zugasti,^{†a} M. R. González-Morales,^a J. Martín-Chivelet^c and J. O. Cáceres^{*b}

The potential application of Mg/Ca ratios in top shells of the mollusc species *Phorcus lineatus* (Da Costa, 1778) obtained by Laser Induced Breakdown Spectroscopy (LIBS) has been evaluated as an environmental proxy to reconstruct paleotemperatures and season of capture of molluscs for the first time. All samples were collected from the Cantabrian Sea (Spain). The results were compared with instrumental sea surface temperatures (SSTs) and with a known reliable proxy as the oxygen isotope ratio ($\delta^{18}\text{O}_{\text{shell}}$) which is mainly dependent on the SST, obtained from the same shells. Measurements were taken in two different biominerals of the shell (aragonite and calcite) resulting in a correlation between Mg/Ca ratios and SSTs of $R^2 = 0.43$ and 0.44 , respectively. Mg/Ca ratios were also studied through a long sequence on three shells collected in autumn 2012. The results show variations in Mg/Ca ratios related to seasonal changes in the SST throughout the year and a good correlation between Mg/Ca ratios and $\delta^{18}\text{O}_{\text{shell}}$ in two shells ($R^2 = 0.70$ and 0.65 , respectively).

Received 8th May 2015
Accepted 15th June 2015

DOI: 10.1039/c5ja00168d

www.rsc.org/jaas

1. Introduction

The reconstruction of the climatic and environmental variability in the past depended on the fundamental task of obtaining high-resolution climate and environmental proxies. As many of these proxies are based on geochemical parameters, this task requires application of efficient analytical techniques to the obtained adequate geological, paleontological, or archaeological records.¹ Biological remains from archaeological sites can provide not only valuable data for reconstructing the past climate but can also yield key information about human activities. Among these, mollusc remains are found in many archaeological sites, commonly as large accumulation of shells as a result of human activities.² These shells can be excellent archives of the environmental conditions that existed during the life of the organism (and changes through it, for example, due to seasonal patterns) as well as indicators of human behaviour (*e.g.*, season of capture of molluscs, subsistence strategies, *etc.*³).

Most advances in the paleoclimatological analysis of these shells are related to the study of stable isotope ratios, mainly

$\delta^{18}\text{O}$, which provide information about the sea water temperature,^{4–8} and much less attention has been paid to other geochemical data of high potential, such as the trace element ratios in the biogenic carbonate. This potential, which is determined by the influence of temperature on the incorporation of chemical elements into the calcium carbonate along the shell growth axis, is however masked by the difficulties in retrieving significant and accurate data from the shells (for example, from micrometric growth layers) and in interpreting the data, as trace element concentrations are also non-environmental factor dependent.

With the aim of progressing in this direction, this paper presents the study of Mg/Ca ratios in the shells of modern specimens of the mollusc *Phorcus lineatus*, a gastropod abundant in present-day coastal areas and in archaeological sites. The relationship between the Mg/Ca ratio of shell carbonate and the temperature of water has been previously explored for several mollusc species under both laboratory and natural conditions, with remarkably different results.

A strong correlation of Mg/Ca with temperature has been shown in experimental work on *Patella rustica* and *Patella caerulea*,⁹ *Pecten maximus*¹⁰ and *Pinna nobilis*.¹¹ However, the correlation is much weaker in the case of *Tridacna gigas*,¹² *Concholepas concholepas*,¹³ *Protothaca staminea*,¹⁴ and *Cassostrea virginica*¹⁵ or almost none in *Mesodesma donacium*,¹⁶ *Chione subrugosa*,¹⁶ *Mytilus californianus*,¹⁷ *Arctica islandica*^{18–20} and *Protothaca thaca*.²¹

This paper evaluates, for the first time, the usage of Laser Induced Breakdown Spectroscopy (LIBS),^{22,23} to obtain the

^aPrehistoric International Research Institute of Cantabria, University of Cantabria, 39005 Santander, Spain^bDepartment of Analytical Chemistry, Faculty of Chemistry, Complutense University, 28040 Madrid, Spain. E-mail: jcaceres@ucm.es; Fax: +34 91 3944329; Tel: +34 91 3944322^cDepartment of Stratigraphy, Faculty of Geological Sciences, Complutense University and IGEO Geosciences Institute (CSIC-UCM), 28040 Madrid, Spain[†] These authors have contributed equally.

Plume dynamics of laser-produced swine muscle tissue plasma

J. J. CAMACHO^{1,*}, L. DIAZ², A. MARIN-ROLDAN³, S. MONCAYO³, and J. O. CACERES³

¹Dpto. de Química-Física Aplicada, Facultad de Ciencias, Universidad Autónoma de Madrid, Cantoblanco, 28049-Madrid, Spain.

²Instituto de Estructura de la Materia, CFMAC, CSIC, Serrano 121, 28006-Madrid, Spain.

³Dpto. de Química Analítica, Facultad de Ciencias Químicas, Universidad Complutense, C. Univ., 28040-Madrid, Spain.

*Corresponding author: j.j.camacho@uam.es

We report on the plume dynamics of the plasma induced by laser ablation of a swine skeletal muscle tissue sample in different vacuum conditions. Pulses from a transversely excited atmospheric CO₂ laser were focused onto a target sample and the induced plasma was allowed to expand in different air pressures. The expansion features were studied using fast photography of the overall visible emission by using a gated intensified charged coupled device. Free expansion and plume splitting were observed at different pressure levels. The expansion of the plasma plume front was analyzed using various expansion models and the velocity of the plume front was estimated. The effect of the number of accumulated laser shots on the crater volume at different ambient air pressures and an elemental analysis of the sample were performed using scanning electron microscopy (SEM) coupled with energy dispersive X-ray spectroscopy (EDS) analysis. The surface morphology of the irradiated surface showed that increasing the pressure of the ambient gas, decreased the ablated mass or in other words it reduced significantly the laser-target coupling.

Index Headings: Laser-produced plasma, LPP; Shock wave; Plasma imaging; Scanning electron microscopy, SEM; Energy dispersive X-ray spectroscopy, EDS.

INTRODUCTION

Laser-produced plasmas (LPPs) are currently a topic of considerable interest in fundamental and applied areas of scientific research. LPPs are transient in nature with characteristic parameters evolving fast in space and time. There are different diagnostic techniques to characterize LPPs such as optical emission spectroscopy, shadowgraphy and fast photography, laser induced fluorescence, Langmuir probing, photo-thermal beam deflection, mass spectrometry, etc. These studies are essential for understanding various processes in fundamental plasma physics. There are many applications for LPPs in a wide variety of fields of basic research and material technology such as fabrication of thin films by pulsed laser deposition,¹ production of nanoparticles and clusters,² surface modifications,³ spectrochemical analysis of various materials through laser-induced breakdown spectroscopy (LIBS),^{4,8} ion source,⁹ lithography,¹⁰ microscopy,¹¹ in the fabrication of micro-electronic devices¹², etc. Beyond traditional applications of LIBS, where inorganic materials are mainly studied, recent progresses in LIBS lead to analysis of biological warfare agent materials,¹³⁻¹⁴ animal tissue studies,¹⁵⁻¹⁶ biological aerosols¹⁷ and identification of bacteria.¹⁸ On the other hand, the CO₂ laser has been extensively used over the past 30 years in various fields of medicine¹⁹ and is still one of the most useful and efficient medical lasers available in the market. This laser is mainly employed a surgical or cosmetic tool and can cut or vaporize tissue with fairly little bleeding and minimal damage to surrounding tissue. It is used to remove thin layers from the surface of the skin without going into the deeper layers.



Technical note

Determination of the postmortem interval by Laser Induced Breakdown Spectroscopy using swine skeletal muscles



A. Marín-Roldan, S. Manzoor, S. Moncayo, F. Navarro-Villoslada, R.C. Izquierdo-Hornillos, J.O. Caceres *

Departamento de Química Analítica, Facultad de Ciencias Químicas, Universidad Complutense, Ciudad Universitaria, 28040 Madrid, Spain

ARTICLE INFO

Article history:

Received 26 November 2012

Accepted 13 July 2013

Available online 20 July 2013

Keywords:

Postmortem interval

Swine

Laser Induced Breakdown Spectroscopy

ABSTRACT

Skin and muscle samples are useful to discriminate individuals as well as their postmortem interval (PMI) in crime scenes and natural or caused disasters. In this study, a simple and fast method based on Laser Induced Breakdown Spectroscopy (LIBS) has been developed to estimate PMI using swine skeletal muscle samples. Environmental conditions (moisture, temperature, fauna, etc.) having strong influence on the PMI determination were considered. Time-dependent changes in the emission intensity ratio for Mg, Na, H α and K were observed, as a result of the variations in their concentration due to chemical reactions in tissues and were correlated with PMI. This relationship, which has not been reported previously in the forensic literature, offers a simple and potentially valuable means of estimating the PMI.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

The postmortem interval (PMI) refers to the time since death and its estimation is extremely important in criminal, civil and forensic investigations [1]. Postmortem decomposition is reported to begin immediately after death involving the breakdown of soft tissues due to the autolysis and putrefaction processes [2], and often with superimposed insect activity [3]. A death event is followed by a number of complicated chemical reactions which continue to occur within the cells, and cause specific time-dependent changes in metabolism and sub-cellular structures, enabling to determine PMI [1]. Many methods have been employed to accurately determine the PMI, also known as time of death (TOD), but the accuracy of these methods still leaves a high time uncertainty and needs improvements [4]. From an entomological point of view, since insect growth and development are heavily influenced by geographical, topological and environmental conditions, the difficulty in accounting for the conditions at the scene prior to discovery introduces some uncertainty in application. This variability in the environmental conditions made a well-characterized foundation of workable information difficult to attain [5]. Actually, methods of PMI estimation include body temperature measurements and postmortem chemical changes in the body [4]. In addition, some methods have also focused on the degradation of nucleic acids including DNA and RNA [1,6–8] and proteins within different tissues [9,10]. However, these methods require specialized staff and laboratories, and in some cases complicated and costly analysis procedures.

A significant relationship has been established between the ratio of postmortem sodium and potassium concentration variation with PMI

in wistar rat serum [11] and human blood [12] samples. However, authors showed that internal and external environmental factors as well as the cause of death significantly affect PMI prediction due to the fast changes of this ratio in the serum and plasma samples (0–12 h) [13]. On the other hand, individual biological variability also limits the usefulness of PMI prediction based on these ratio measurements. Ideally, the concentration of the constituents at any given postmortem interval should display minimal individual biological variation and should be independent of environmental conditions.

Laser Induced Breakdown Spectroscopy (LIBS) has been a subject of research for the past few decades because of its unique features and wide variety of applications in various fields [14,15]. In recent years, LIBS has become a powerful analytical tool because of its ability to carry out a rapid qualitative and quantitative analysis of different samples [16,17]. LIBS analyzes a sample by direct measurement of the atomic emission of the elements from laser-induced plasma generated by the ablation of the sample, providing an immediate spectral fingerprint which is representative of its elemental composition [18]. Moreover, the technique requires little or no sample preparation as opposed to biochemical methods.

The spatial distribution of volumetric energy density generated by laser irradiation drives all pulsed laser ablation processes [19]. In case of laser–tissue interaction, this distribution is controlled by the incident radiant exposure and the optical absorption and scattering properties of the tissue. A detailed description of laser–tissue interaction has been the subject of many reviews discussing general aspects of this process [20–22]. Specific laser ablation processes resulting from nonlinear absorption have also been considered [23,24].

This paper focuses on the development of a simple, fast and direct analytical method for the determination of PMI based on LIBS using muscle tissue samples. Swine muscle tissues were selected due to similar

* Corresponding author.

E-mail address: jcaceres@quim.ucm.es (J.O. Caceres).



Rapid identification and discrimination of bacterial strains by laser induced breakdown spectroscopy and neural networks

S. Manzoor^a, S. Moncayo^a, F. Navarro-Villoslada^a, J.A. Ayala^b, R. Izquierdo-Hornillos^a, F.J. Manuel de Villena^a, J.O. Cáceres^{a,*}

^a Departamento de Química Analítica, Facultad de Ciencias Químicas, Universidad Complutense, 28040 Madrid, Spain

^b Centro de Biología Molecular "Severo Ochoa", CSIC. C/Nicolás Cabrera, 1, Cantoblanco, 28049 Madrid, Spain

ARTICLE INFO

Article history:

Received 18 October 2013

Received in revised form

16 December 2013

Accepted 24 December 2013

Available online 4 January 2014

Keywords:

Bacterial strain discrimination

Laser induced breakdown spectroscopy

(LIBS)

Neural networks (NN)

Bacteria

Hospital acquired infections (HAI)

ABSTRACT

Identification and discrimination of bacterial strains of same species exhibiting resistance to antibiotics using laser induced breakdown spectroscopy (LIBS) and neural networks (NN) algorithm is reported. The method has been applied to identify 40 bacterial strains causing hospital acquired infections (HAI), i.e. *Pseudomonas aeruginosa*, *Escherichia coli*, *Klebsiella pneumoniae*, *Salmonella typhimurium*, *Salmonella pullorum* and *Salmonella salamae*. The strains analyzed included both isolated from clinical samples and constructed in laboratory that differ in mutations as a result of their resistance to one or more antibiotics. Small changes in the atomic composition of the bacterial strains, as a result of their mutations and genetic variations, were detected by the LIBS–NN methodology and led to their identification and classification. This is of utmost importance because solely identification of bacterial species is not sufficient for disease diagnosis and identification of the actual strain is also required. The proposed method was successfully able to discriminate strains of the same bacterial species. The optimized NN models provided reliable bacterial strain identification with an index of spectral correlation higher than 95% for the samples analyzed, showing the potential and effectiveness of the method to address the safety and social-cost HAI-related issue.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Hospital acquired infections (HAI) have been widely studied in the past 30 years and have been raised to top-priority issue due to the associated economic and social costs [1]. Therefore, many preventive campaigns as well as new protocols have been implemented [2,3]. On average 5–7% of hospitalized patients are affected by HAI, and 1% of such unwanted events result in the patient's death [4]. Bacteria are responsible for 95% of HAI, *Escherichia coli* (18.2%), *Staphylococcus* (18.1%), *Pseudomonas* (6.0%), *Enterococcus* (15.4%), *Klebsiella* (3.7%), *Acinetobacter* (0.8%), and *Salmonella* (2.8%) being the most relevant ones.

An important issue highlighted in recent years has been the increasing emergence of bacteria that are resistant to many antimicrobial therapies, sometimes resulting in multidrug-resistant strains or "super bugs." One of the overriding reasons for this is the widespread indiscriminate use of antibiotics to treat infections [5]. This antibiotic resistance is evolved under the treatment regimens of single or multidrug combinations as a result of the mutations [6].

During the past decades several methods have been proposed to optimize the identification of bacterial strain, which are based on molecular techniques such as fluorescent probes [7], microarray assemblies [8,9] and polymerase chain reaction [10,11]. However, these methodologies present some difficulties and drawbacks such as use of consumables, primer, probes or fluorescently labeled RNA antibodies [12]. Moreover, sometimes the sequences in the database are not accurate or up-to-date and micro-heterogeneity is also found common in 16S rRNA gene sequence within a species [13,14]. The phenotypic similarities between the strains of the same bacterial species restrict their identification using routine diagnostic methods [15]. Although these methods provide a reliable and accurate bacterial identification, special sample treatment methods, the high costs and low speed to perform such analysis limit their use as rapid diagnostic methods in common laboratories in order to provide quick results which leads to an increase in the rates of infectious diseases in clinical settings. Further the direct handling of these potentially pathogenic bacterial samples poses health-associated security risks. At present, clinical safety procedures and cost-related considerations do not allow an easy routine analysis of highly dangerous pathogenic bacterial specimens causing hospital acquired infections. Nevertheless, bacterial identification within the first 24 h of infection allows the use of a more effective and less risky targeted-therapy decreasing unnecessary hospitalization days and costs.

* Corresponding author. Tel.: +34 91 3944322; fax: +34 91 3944329.
E-mail address: jcaceres@ucm.es (J.O. Cáceres).